

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

PATENT COOPERATION TREATY

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

From the INTERNATIONAL BUREAU

To:

United States Patent and Trademark
Office
(Box PCT)
Crystal Plaza 2
Washington, DC 20231
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing: 04 September 1997 (04.09.97)	
International application No.: PCT/AU97/00124	Applicant's or agent's file reference:
International filing date: 28 February 1997 (28.02.97)	Priority date: 01 March 1996 (01.03.96)
Applicant: BRUGLIERA, Filippa et al	

1. The designated Office is hereby notified of its election made:

☒ in the demand filed with the International preliminary Examining Authority on:
14 July 1997 (14.07.97)☐ in a notice effecting later election filed with the International Bureau on:2. The election ☒ was☐ was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland Facsimile No.: (41-22) 740.14.35	Authorized officer: J. Zahra Telephone No.: (41-22) 338.83.38
---	---

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference PN8386/EJH/AF	FOR FURTHER ACTION see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.	
International application No. PCT/AU 97/00124	International filing date (<i>day/month/year</i>) 28 February 1997	(Earliest) Priority Date (<i>day/month/year</i>) 1 March 1996
Applicant (1) Florigene Limited (2) Brugliera F., et al.		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of **3** sheets.

☒ It is also accompanied by a copy of each prior art document cited in this report.

1. ☐ Certain claims were found unsearchable (See Box I)
2. ☐ Unity of invention is lacking (See Box II)
3. ☒ The international application contains disclosure of a **nucleotide and/or amino acid sequence listing** and the international search was carried out on the basis of the sequence listing

☒ filed with the international application
☐ furnished by the applicant separately from the international application,

☐ but not accompanied by a statement to the effect that it did not include matter going beyond the disclosure in the international application as filed

☐ transcribed by this Authority
4. With regard to the **title**, ☒ the text is approved as submitted by the applicant.
☐ the text has been established by this Authority to read as follows:
5. With regard to the **abstract**, ☒ the text is approved as submitted by the applicant
☐ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.
6. The figure of the **drawings** to be published with the abstract is:

Figure No.

☐ as suggested by the applicant.
☐ because the applicant failed to suggest a figure
☐ because this figure better characterises the invention
☒ None of the figures

A. CLASSIFICATION OF SUBJECT MATTERInt Cl⁶: C12N 15/29, 15/63
A01H 5/00, 5/02**B. FIELDS SEARCHED**Minimum documentation searched (classification system followed by classification symbols)
See electronic databases belowDocumentation searched other than minimum documentation to the extent that such documents are included in the fields searched
See electronic databases belowElectronic data base consulted during the international search (name of data base and, where practicable, search terms used)
WPAT, JAPIO, USPM - ss1: flavonoid(3n)monooxygenase# or flavonoid(3n)hydroxylase# or F3(OH); ss2: A01H/IC or C12N-009/IC or C12N-015/IC; ss1 and ss2
CHEMICAL ABSTRACTS - flavonoid 3' hydroxylase (clones pcgp772, pcgp602, pcgp619, pcgp635, pcgp854, pcgp773)/cn; flavonoid 3' hydroxylase/cn, flavonoid 3' monooxygenase
GENEBANK, EMBL, SWISS-PROT - Sequence searches**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	AU A 37413/93 (INTERNATIONAL FLOWER DEVELOPMENTS PTY. LTD.) 14 OCTOBER 1993 C12N 15/53, 15/11, 9/02, C12Q 1/66, A01H 5/00	27,28,30,31,38
A	PHYTOCHEMISTRY, VOL. 35 PAGES 145-150, 1994, SCHWINN KE et al. "Floral flavonoids and the potential for pelargonidin biosynthesis in commercial chrysanthemum cultivars"	1 - 38

☐ Further documents are listed in the continuation of Box C☒ See patent family annex

* Special categories of cited documents:	
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E" earlier document but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
9 April 1997Date of mailing of the international search report
22 APR 1997Name and mailing address of the ISA/AU
AUSTRALIAN INDUSTRIAL PROPERTY ORGANISATION
PO BOX 200
WODEN ACT 2606
AUSTRALIA Facsimile No.: (06) 285 3929Authorized officer
KAREN AYERS
Telephone No.: (06) 283 2082

INTERNATIONAL SEARCH REPORT
Information on patent family members

International Application No.
PCT/AU 97/00124

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report	Patent Family Member
AU, A, 37413/93	WO, A, 93/20206
	EP, A, 640136
	NZ, A, 249808
END OF ANNEX	

Applicant's or agent's file reference PN8386/EJH/AF	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416).
International application No. PCT/AU 97/00124	International filing date 28 February 1997	Priority Date 1 March 1996
International Patent Classification (IPC) or national classification and IPC Int. Cl.⁶ C12N 15/29, 15/52, 15/82; A01H 4/00, 5/00, 5/02, 5/08, 5/10, 5/12		
Applicant (1) FLORIGENE LIMITED (2) BRUGLIERA F., ET AL.		

1.	This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
2.	This REPORT consists of a total of 4 sheets, including this cover sheet. <input type="checkbox"/> This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT). These annexes consist of a total of sheet(s).
3.	This report contains indications relating to the following items: I <input checked="" type="checkbox"/> Basis of the report II <input type="checkbox"/> Priority III <input type="checkbox"/> Non-establishment of opinion with regard to novelty, inventive step and industrial applicability IV <input checked="" type="checkbox"/> Lack of unity of invention V <input checked="" type="checkbox"/> Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement VI <input type="checkbox"/> Certain documents cited VII <input type="checkbox"/> Certain defects in the international application VIII <input type="checkbox"/> Certain observations on the international application

Date of submission of the demand 14 July 1997	Date of completion of the report 24 October 1997
Name and mailing address of the IPEA/AU AUSTRALIAN INDUSTRIAL PROPERTY ORGANISATION PO BOX 200 WODEN ACT 2606 AUSTRALIA Facsimile No. (02) 6285 3929	Authorized Officer KAREN AYERS Telephone No. (02) 6283 2082

I. Basis of the report

1. This report has been drawn on the basis of *(Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.)*:

☒ the international application as originally filed.

☐ the description, pages , as originally filed,
 pages , filed with the demand,
 pages , filed with the letter of ,
 pages , filed with the letter of .

☐ the claims, Nos. , as originally filed,
 Nos. , as amended under Article 19,
 Nos. , filed with the demand,
 Nos. , filed with the letter of ,
 Nos. , filed with the letter of .

☐ the drawings, sheets/fig , as originally filed,
 sheets/fig , filed with the demand,
 sheets/fig , filed with the letter of ,
 sheets/fig , filed with the letter of .

2. The amendments have resulted in the cancellation of:

☐ the description, pages

☐ the claims, Nos.

☐ the drawings, sheets/fig

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).

4. Additional observations, if necessary:

IV. Lack of unity of invention

1. In response to the invitation to restrict or pay additional fees the applicant has:

- ☐ restricted the claims.
- ☐ paid additional fees.
- ☐ paid additional fees under protest.
- ☐ neither restricted nor paid additional fees.

2. ☐ This Authority found that the requirement of unity of invention is not complied with and chose, according to Rule 68.1, not to invite the applicant to restrict or pay additional fees.

3. This Authority considers that the requirement of unity of invention in accordance with Rules 13.1, 13.2 and 13.3 is

- ☒ complied with.
- ☐ not complied with for the following reasons:

4. Consequently, the following parts of the international application were the subject of international preliminary examination in establishing this report:

- ☒ all parts.
- ☐ the parts relating to claims Nos.

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**1. Statement**

Novelty (N)	Claims 1-26, 29, 32-37	YES
	Claims 27, 28, 30, 31, 38	NO
Inventive step (IS)	Claims 1-26, 29, 32-37	YES
	Claims 27, 28, 30, 31, 38	NO
Industrial applicability (IA)	Claims 1-38	YES
	Claims	NO

2. Citations and explanations**Citations**

D1: AU A 37413/93 (International Flower Developments Pty. Ltd.) 14 October 1993 (14.10.93) C12N 15/53, 15/11, 9/02, C12Q 1/66, A01H 5/00

D2: Phytochemistry volume 35, pages 145-150 (1994) by Schwinn KE et al. "Floral flavonoids and the potential for pelargonidin biosynthesis in commercial chrysanthemum cultivars".

D1 is considered to be the closest related art and discloses nucleotide and amino acid sequences encoding a flavonoid-3'-hydroxylase. Claims 1 to 26, 29 and 32 to 37 are restricted to specific sequences encoding flavonoid-3'-hydroxylase which are exemplified in the description as SEQ. ID. Nos: 1-26. However, none of these sequences are disclosed in D1, and therefore claims 1 to 26, 29 and 32 to 37 are novel and inventive in light of D1.

Claims 27, 28, 30, 31 and 38 are not limited to specific sequences exemplified in the description, so are not considered novel or inventive when compared to D1 which discloses a flavonoid-3'-hydroxylase.

D2 defines general information on the role of flavonoid-3'-hydroxylase in flavonoid production and pigmentation of chrysanthemums. No nucleotide or amino acid sequences have been identified and therefore claims 1 to 32 are novel and inventive in light of this document.

Claims 1 to 38 possess industrial applicability for the manipulation of flower colour in the commercial flower industry.

142,108
PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

International Patent Classification⁶:

C12N 15/29, 15/63, A01H 5/00, 5/02

A1

(11) International Publication Number:

WO 97/32023

(43) International Publication Date:

4 September 1997 (04.09.97)

(21) International Application Number: PCT/AU97/00124

(22) International Filing Date: 28 February 1997 (28.02.97)

(30) Priority Data:

PN 8386

1 March 1996 (01.03.96)

AU

(71) Applicant (for all designated States except US): FLORIGENE LIMITED [AU/AU]; 16 Gipps Street, Collingwood, VIC 3066 (AU).

(72) Inventors; and

(75) Inventors/Applicants (for US only): BRUGLIERA, Filippa [AU/AU]; 11 Kalimna Street, Preston, VIC 3072 (AU). HOLTON, Timothy, Albert [AU/AU]; Unit 1, 8 May Street, Elwood, VIC 3184 (AU). MICHAEL, Michael, Zenon [AU/AU]; 15 Goodwin Street, West Preston, VIC 3072 (AU).

(74) Agents: HUGHES, E., John, L. et al.; Davies Collison Cave, 1 Little Collins Street, Melbourne, VIC 3000 (AU).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

(54) Title: GENETIC SEQUENCES ENCODING FLAVONOID PATHWAY ENZYMES AND USES THEREFOR

(57) Abstract

The present invention relates generally to genetic sequences encoding flavonoid pathway metabolising enzymes and more particularly to flavonoid 3'-hydroxylase (hereinafter referred to as "F3'H") or derivatives thereof and their use in the manipulation of pigmentation in plants and other organisms.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

GENETIC SEQUENCES ENCODING FLAVONOID PATHWAY ENZYMES AND USES THEREFOR

The present invention relates generally to genetic sequences encoding flavonoid pathway
5 metabolising enzymes and more particularly to flavonoid 3'-hydroxylase (hereinafter referred
to as "F3'H") or derivatives thereof and their use in the manipulation of pigmentation in
plants and other organisms.

Bibliographic details of the publications referred to by the author in this specification are
10 collected at the end of the description. Sequence Identity Numbers (SEQ ID NOs) for the
nucleotide and amino acid sequences referred to in the specification and claims are defined
following the bibliography. A summary of the SEQ ID NOs, and the sequences to which
they relate, is provided prior to the Examples.

15 Throughout this specification, unless the context requires otherwise, the word "comprise",
or variations such as "comprises" or "comprising", will be understood to imply the inclusion
of a stated element or integer or group of elements or integers but not the exclusion of any
other element or integer or group of elements or integers.

20 The rapidly developing sophistication of recombinant DNA technology is greatly facilitating
research and development in a range of biotechnology related industries. The horticultural
industry has become a recent beneficiary of this technology which has contributed to
developments in disease resistance in plants and flowers exhibiting delayed senescence after
cutting. Some attention has also been directed to manipulating flower colour.

25

The flower industry strives to develop new and different varieties of flowering plants. An
effective way to create such novel varieties is through the manipulation of flower colour.
Classical breeding techniques have been used with some success to produce a wide range of
colours for most of the commercial varieties of flowers. This approach has been limited,
30 however, by the constraints of a particular species' gene pool and for this reason it is rare for

- 2 -

a single species to have a full spectrum of coloured varieties. In addition, traditional breeding techniques lack precision. The aesthetic appeal of the flower is a combination of many factors such as form, scent and colour; modification of one character through hybridization can often be at the expense of an equally valuable feature. The ability to genetically engineer precise colour changes in cutflower and ornamental species would offer significant commercial opportunities in an industry which has rapid product turnover and where novelty is an important market characteristic.

Flower colour is predominantly due to two types of pigment: flavonoids and carotenoids. Flavonoids contribute to a range of colours from yellow to red to blue. Carotenoids impart an orange or yellow tinge and are commonly the major pigment in yellow or orange flowers. The flavonoid molecules which make the major contribution to flower colour are the anthocyanins which are glycosylated derivatives of cyanidin, delphinidin, petunidin, peonidin, malvidin and pelargonidin, and are localised in the vacuole. The different anthocyanins can produce marked differences in colour. Flower colour is also influenced by co-pigmentation with colourless flavonoids, metal complexation, glycosylation, acylation and vacuolar pH (Forkmann, 1991).

The biosynthetic pathway for the flavonoid pigments (hereinafter referred to as the "flavonoid pathway") is well established and is shown in Figures 1a and 1b (Ebel and Hahlbrock, 1988; Hahlbrock and Grisebach, 1979; Wiering and De Vlaming, 1984; Schram *et al.*, 1984; Stafford, 1990; Van Tunen and Mol, 1990; Dooner *et al.*, 1991; Martin and Gerats, 1993; Holton and Cornish, 1995). The first committed step in the pathway involves the condensation of three molecules of malonyl-CoA with one molecule of *p*-coumaroyl-CoA. This reaction is catalysed by the enzyme chalcone synthase (CHS). The product of this reaction, 2',4,4',6', tetrahydroxy-chalcone, is normally rapidly isomerized to produce naringenin by the enzyme chalcone flavanone isomerase (CHI). Naringenin is subsequently hydroxylated at the 3 position of the central ring by flavanone 3-hydroxylase (F3H) to produce dihydrokaempferol (DHK).

- 3 -

The pattern of hydroxylation of the B-ring of DHK plays a key role in determining petal colour. The B-ring can be hydroxylated at either the 3', or both the 3' and 5' positions, to produce dihydroquercetin (DHQ) and dihydromyricetin (DHM), respectively. Two key enzymes involved in this pathway are flavonoid 3'-hydroxylase and flavonoid 3',5'-hydroxylase, both of the cytochrome P450 class. Cytochrome P450 enzymes are widespread in nature and genes have been isolated and sequenced from vertebrates, insects, yeasts, fungi, bacteria and plants.

Flavonoid 3'-hydroxylase acts on DHK to produce DHQ and on naringenin to produce eriodictyol. Reduction and glycosylation of DHQ produces the cyanidin-glycoside and peonidin-glycoside pigments which, in many plant species (for example rose, carnation and chrysanthemum), contribute to red and pink flower colour. The synthesis of these anthocyanins can also result in other flower colours. For example, blue cornflowers contain cyanin. The ability to control flavonoid 3'-hydroxylase activity, or other enzymes involved in the flavonoid pathway, in flowering plants would provide a means to manipulate petal colour. Different coloured versions of a single cultivar could thereby be generated and in some instances a single species would be able to produce a broader spectrum of colours.

A nucleotide sequence (referred to herein as SEQ ID NO:26) encoding a petunia flavonoid 3'-hydroxylase has been cloned (see International Patent Application No. PCT/AU93/00127 [WO 93/20206]). However, this sequence was inefficient in its ability to modulate the production of 3'-hydroxylated anthocyanins in plants. There is a need, therefore, to identify further genetic sequences encoding flavonoid 3'-hydroxylases which efficiently modulate the hydroxylation of flavonoid compounds in plants. More particularly, there is a need to identify further genetic sequences encoding flavonoid 3'-hydroxylases which efficiently modulate the production of 3'-hydroxylated anthocyanins in plants.

In accordance with the present invention, genetic sequences encoding flavonoid 3'-hydroxylase have been identified and cloned. The recombinant genetic sequences of the present invention permit the modulation of expression of genes encoding this enzyme by, for

- 4 -

example, *de novo* expression, over-expression, suppression, antisense inhibition and ribozyme activity. The ability to control flavonoid 3'-hydroxylase synthesis in plants permits modulation of the composition of individual anthocyanins as well as alteration of relative levels of flavonols and anthocyanins, thereby enabling the manipulation of tissue colour, such as petals, leaves, seeds and fruit. The present invention is hereinafter described in relation to the manipulation of flower colour but this is done with the understanding that it extends to manipulation of other plant tissues, such as leaves, seeds and fruit.

Accordingly, one aspect of the present invention provides an isolated nucleic acid molecule comprising a sequence of nucleotides encoding a flavonoid 3'-hydroxylase or a derivative thereof wherein said flavonoid 3'-hydroxylase or its derivative is capable of more efficient modulation of hydroxylation of flavonoid compounds in plants than is a flavonoid 3'-hydroxylase encoded by the nucleotide sequence set forth in SEQ ID NO:26.

Efficiency as used herein relates to the capability of the flavonoid 3'-hydroxylase enzyme to hydroxylate flavonoid compounds in a plant cell. This provides the plant with additional substrates for other enzymes of the flavonoid pathway able to further modify this molecule, via, for example, glycosylation, acylation and rhamnosylation, to produce various anthocyanins which contribute to the production of a range of colours. The modulation of 3'-hydroxylated anthocyanins is thereby permitted. Efficiency is conveniently assessed by one or more parameters selected from: extent of transcription, as determined by the amount of mRNA produced; extent of hydroxylation of naringenin and/or DHK; extent of translation of mRNA, as determined by the amount of translation product produced; extent of production of anthocyanin derivatives of DHQ or DHM; the extent of effect on tissue colour, such as flowers, seeds, leaves or fruits.

Another aspect of the present invention is directed to an isolated nucleic acid molecule comprising a sequence of nucleotides which maps to the genetic locus designated *Ht1* or *Ht2* in petunia, or to equivalent such loci in other flowering plant species, and wherein said isolated nucleic acid molecule encodes, or is complementary to a sequence which encodes,

- 5 -

a flavonoid 3'-hydroxylase.

A further aspect of the present invention contemplates an isolated nucleic acid molecule comprising a sequence of nucleotides which corresponds to the genetic locus designated Ht1 5 or Ht2 in petunia, or to loci in other flowering plant species which contain sequences which control production of 3'-hydroxylated flavonoids, and wherein said isolated nucleic acid molecule encodes a flavonoid 3'-hydroxylase or a derivative thereof which is capable of more efficient conversion of DHK to DHQ in plants than is the flavonoid 3'-hydroxylase set forth in SEQ ID NO:26.

10

In accordance with the above aspects of the present invention there is provided a nucleic acid molecule comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:1 or having at least about 60% similarity thereto or capable of hybridising to the sequence set forth in SEQ ID NO:1 under low stringency 15 conditions.

In a related embodiment, there is provided a nucleic acid molecule comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:3 or having at least about 60% similarity thereto or capable of hybridising to the sequence set 20 forth in SEQ ID NO:3 under low stringency conditions.

In another related embodiment, the present invention is directed to a nucleic acid molecule comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:5 or having at least about 60% similarity thereto or capable of 25 hybridising to the sequence set forth in SEQ ID NO:5 under low stringency conditions.

Yet another related embodiment provides a nucleic acid molecule comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:7 or having at least about 60% similarity thereto or capable of hybridising to the sequence set 30 forth in SEQ ID NO:7 under low stringency conditions.

- 6 -

Still yet a further embodiment of the present invention relates to a nucleic acid molecule comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:9 or having at least about 60% similarity to the coding region thereof or capable of hybridising to the sequence set forth in SEQ ID NO:9 under low stringency
5 conditions.

In another further embodiment, there is provided a nucleic acid molecule comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:14 or having at least about 60% similarity thereto or capable of hybridising to the
10 sequence set forth in SEQ ID NO:14 under low stringency conditions.

In yet another further embodiment, the present invention is directed to a nucleic acid molecule comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:16 or having at least about 60% similarity thereto or capable of
15 hybridising to the sequence set forth in SEQ ID NO:16 under low stringency conditions.

Still yet another further embodiment provides a nucleic acid molecule comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:18 or having at least about 60% similarity thereto or capable of hybridising to the sequence set
20 forth in SEQ ID NO:18 under low stringency conditions.

Moreover, yet a further embodiment of the present invention relates to a nucleic acid molecule comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:20 or having at least about 60% similarity thereto or
25 capable of hybridising to the sequence set forth in SEQ ID NO:20 under low stringency conditions.

Still yet another further embodiment is directed to a nucleic acid molecule comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:22 or having at least about 60% similarity thereto or capable of hybridising to the
30

- 7 -

sequence set forth in SEQ ID NO:22 under low stringency conditions.

In still yet another further embodiment, the present invention provides a nucleic acid molecule comprising a nucleotide sequence or complementary nucleotide sequence substantially as set
5 forth in SEQ ID NO:24 or having at least about 60% similarity thereto or capable of hybridising to the sequence set forth in SEQ ID NO:24 under low stringency conditions.

In a particularly preferred embodiment there is provided an isolated nucleic acid molecule comprising a nucleotide sequence or complementary nucleotide sequence substantially as set
10 forth in SEQ ID NO:1 or having at least about 60% similarity thereto or capable of hybridising to the sequence set forth in SEQ ID NO:1 under low stringency conditions, wherein said nucleotide sequence maps to the genetic locus designated Ht1 or Ht2 in petunia, or to equivalent such loci in other flowering plant species, and wherein said isolated nucleic acid molecule encodes, or is complementary to a sequence which encodes, a flavonoid 3'-
15 hydroxylase.

Reference herein to a low stringency at 42°C includes and encompasses from at least about 1% to at least about 15% formamide and from at least about 1M to at least about 2M salt for hybridization, and at least about 1M to at least about 2M salt for washing conditions.
20 Alternative stringency conditions may be applied where necessary, such as medium stringency, which includes and encompasses from at least about 16% to at least about 30% formamide and from at least about 0.5M to at least about 0.9M salt for hybridization, and at least about 0.5M to at least about 0.9M salt for washing conditions, or high stringency, which includes and encompasses from at least about 31% to at least about 50% formamide and from
25 at least about 0.01M to at least about 0.15M salt for hybridization, and at least about 0.01M to at least about 0.15M salt for washing conditions. Hybridization may be carried out at different temperatures and, where this occurs, other conditions may be adjusted accordingly.

Another aspect of the present invention provides a nucleic acid molecule comprising a
30 sequence of nucleotides encoding or complementary to a sequence encoding an amino acid

- 8 -

sequence substantially as set forth in SEQ ID NO:2 or an amino acid sequence having at least about 50% similarity thereto.

In a related embodiment, there is provided a nucleic acid molecule comprising a sequence of
5 nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:4 or an amino acid sequence having at least about 50% similarity thereto.

A further related embodiment of the present invention is directed to a nucleic acid molecule
10 comprising a sequence of nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:6 or an amino acid sequence having at least about 50% similarity thereto.

Still another related embodiment provides a nucleic acid molecule comprising a sequence of
15 nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:8 or an amino acid sequence having at least about 50% similarity thereto.

Yet still another related embodiment relates to a nucleic acid molecule comprising a sequence
20 of nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:10 or SEQ ID NO:11 or SEQ ID NO:12 or SEQ ID NO:13 or an amino acid sequence having at least about 50% similarity thereto.

In another further embodiment, there is provided a nucleic acid molecule comprising a
25 sequence of nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:15 or an amino acid sequence having at least about 50% similarity thereto.

In yet another further embodiment, the present invention is directed to a nucleic acid molecule
30 comprising a sequence of nucleotides encoding or complementary to a sequence encoding an

- 9 -

amino acid sequence substantially as set forth in SEQ ID NO:17 or an amino acid sequence having at least about 50% similarity thereto.

Still yet another further embodiment provides a nucleic acid molecule comprising a sequence
5 of nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:19 or an amino acid sequence having at least about 50% similarity thereto.

Moreover, yet a further embodiment of the present invention relates to a nucleic acid
10 molecule comprising a sequence of nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:21 or an amino acid sequence having at least about 50% similarity thereto.

Still yet another further embodiment is directed to a nucleic acid molecule comprising a
15 sequence of nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:23 or an amino acid sequence having at least about 50% similarity thereto.

In still yet another further embodiment, the present invention provides a nucleic acid molecule
20 comprising a sequence of nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:25 or an amino acid sequence having at least about 50% similarity thereto.

In a particularly preferred embodiment there is provided an isolated nucleic acid molecule
25 comprising a sequence of nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:2 or an amino acid sequence having at least about 50% similarity thereto, wherein said sequence of nucleotides maps to the genetic locus designated Ht1 or Ht2 in petunia, or to equivalent such loci in other flowering plant species, and wherein said isolated nucleic acid molecule encodes, or is
30 complementary to a sequence which encodes, a flavonoid 3'-hydroxylase or a derivative

- 10 -

therof.

The term "similarity" as used herein includes exact identity between compared sequences, at the nucleotide or amino acid level. Where there is non-identity at the nucleotide level, "similarity" includes differences between sequences which result in different amino acids that are nevertheless related to each other at the structural, functional, biochemical and/or conformational levels. Where there is non-identity at the amino acid level, "similarity" includes amino acids that are nevertheless related to each other at the structural, functional, biochemical and/or conformational levels.

10

The nucleic acid molecule defined by SEQ ID NO:1 encodes a flavonoid 3'-hydroxylase from petunia. Examples of other suitable F3'H genes are from carnation (SEQ ID NO:3), snapdragon (SEQ ID NO:5), arabidopsis (SEQ ID NO:7), arabidopsis genomic DNA clone (SEQ ID NO: 9), rose (SEQ ID NO:14), chrysanthemum (SEQ ID NO:16), torenia (SEQ ID NO:18), Japanese morning glory (SEQ ID NO:20), gentian (SEQ ID NO:22) and lisianthus (SEQ ID NO:24). Although the present invention is particularly exemplified by the aforementioned F3'H genes, the subject invention extends to F3'H genes from any species of plant provided that the F3'H gene has at least about 60% similarity at the nucleotide level to a nucleic acid molecule selected from SEQ ID NO:1 or 3 or 5 or 7 or 14 or 16 or 18 or 20 or 22 or 24, or at least about 50% similarity at the amino acid level to an amino acid molecule selected from SEQ ID NO: 2 or 4 or 6 or 8 or 10, 11, 12, 13 or 15 or 17 or 19 or 21 or 23 or 25. The subject invention further extends to F3'H genes from any species of plant provided that the F3'H gene has at least about 60% similarity at the nucleotide level to the coding region of SEQ ID NO:9.

25

The nucleic acid molecules of the present invention are generally genetic sequences in a non-naturally-occurring condition. Generally, this means isolated away from its natural state or synthesized or derived in a non-naturally-occurring environment. More specifically, it includes nucleic acid molecules formed or maintained *in vitro*, including genomic DNA fragments, recombinant or synthetic molecules and nucleic acids in combination with

30

- 11 -

heterologous nucleic acids. It also extends to the genomic DNA or cDNA or part thereof encoding F3'H or part thereof in reverse orientation relative to its or another promoter. It further extends to naturally-occurring sequences following at least a partial purification relative to other nucleic acid sequences.

5

The term "nucleic acid molecule" includes a nucleic acid isolate and a genetic sequence and is used herein in its most general sense and encompasses any contiguous series of nucleotide bases specifying directly, or *via* a complementary series of bases, a sequence of amino acids in a F3'H. Such a sequence of amino acids may constitute a full-length F3'H or an active
10 truncated form thereof or may correspond to a particular region such as an N-terminal, C-terminal or internal portion of the enzyme. The nucleic acid molecules contemplated herein also encompass oligonucleotides useful as genetic probes or as "antisense" molecules capable of regulating expression of the corresponding gene in a plant. An "antisense molecule" as used herein may also encompass a gene construct comprising the structural genomic or cDNA
15 gene or part thereof in reverse orientation relative to its own or another promoter. Accordingly, the nucleic acid molecules of the present invention may be suitable for use as cosuppression molecules, ribozyme molecules, sense molecules and antisense molecules to modulate levels of 3'-hydroxylated anthocyanins.

20 In one embodiment, the nucleic acid molecule encoding F3'H or various derivatives thereof is used to reduce the activity of an endogenous F3'H, or alternatively the nucleic acid molecule encoding this enzyme or various derivatives thereof is used in the antisense orientation to reduce activity of the F3'H. Although not wishing to limit the present invention to any one theory, it is possible that the introduction of the nucleic acid molecule
25 into a cell results in this outcome either by decreasing transcription of the homologous endogenous gene or by increasing turnover of the corresponding mRNA. This may be achieved using gene constructs containing F3'H nucleic acid molecules or various derivatives thereof in either the sense or the antisense orientation. In a further alternative, ribozymes could be used to inactivate target nucleic acid molecules. Alternatively, the nucleic acid
30 molecule encodes a functional F3'H and this is used to elevate levels of this enzyme in plants.

- 12 -

Reference herein to the altering of flavonoid F3'H activity relates to an elevation or reduction in activity of up to 30% or more preferably of 30-50%, or even more preferably 50-75% or still more preferably 75% or greater above or below the normal endogenous or existing levels of activity. The level of activity can be readily assayed using a modified version of the method described by Stotz and Forkmann (1982) (see Example 7) or by assaying for the amount of F3'H product such as quercetin, cyanidin or peonidin as set forth in Example 5.

The present invention further extends to nucleic acid molecules in the form of oligonucleotide primers or probes capable of hybridizing to a portion of the nucleic acid molecules contemplated above, and in particular those selected from the nucleic acid molecules set forth in SEQ ID NOs: 1, 3, 5, 7, 9, 14, 16, 18, 20, 22 or 24 under high, preferably under medium and most preferably under low stringency conditions. Preferably the portion corresponds to the 5' or the 3' end of the F3'H gene. For convenience the 5' end is considered herein to define a region substantially between the 5' end of the primary transcript to a centre portion of the gene, and the 3' end is considered herein to define a region substantially between the centre portion of the gene and the 3' end of the primary transcript. It is clear, therefore, that oligonucleotides or probes may hybridize to the 5' end or the 3' end or to a region common to both the 5' and the 3' ends.

The nucleic acid molecule or its complementary form may encode the full-length enzyme or a part or derivative thereof. By "derivative" is meant any single or multiple amino acid substitutions, deletions, and/or additions relative to the naturally-occurring enzyme and includes parts, fragments, portions, fusion molecules, homologues and analogues. In this regard, the nucleic acid includes the naturally-occurring nucleotide sequence encoding F3'H or may contain single or multiple nucleotide substitutions, deletions and/or additions to said naturally-occurring sequence. A fusion molecule may be a fusion between nucleotide sequences encoding two or more F3'Hs, or a fusion between a nucleotide sequence encoding an F3'H and a nucleotide sequence encoding any other proteinaceous molecule. Fusion molecules are useful in altering substrate specificity.

- 13 -

A derivative of the nucleic acid molecule or its complementary form, or of a F3'H, of the present invention may also include a "part", whether active or inactive. An active or functional nucleic acid molecule is one which encodes an enzyme with F3'H activity. An active or functional molecule further encompasses a partially-active molecule; for example, 5 an F3'H with reduced substrate specificity would be regarded as partially active. A derivative of a nucleic acid molecule may be useful as an oligonucleotide probe, as a primer for polymerase chain reactions or in various mutagenic techniques, for the generation of antisense molecules or in the construction of ribozymes. They may also be useful in developing co-suppression constructs. The nucleic acid molecule according to this aspect of the present 10 invention may or may not encode a functional F3'H. A "part" may be derived from the 5' end or the 3' end or a region common to both the 5' and the 3' ends of the nucleic acid molecule.

Amino acid insertional derivatives of the F3'H of the present invention include amino and/or 15 carboxyl terminal fusions as well as intra-sequence insertions of single or multiple amino acids. Insertional amino acid sequence variants are those in which one or more amino acid residues are introduced into a predetermined site in the protein although random insertion is also possible with suitable screening of the resulting product. Deletional variants are characterised by the removal of one or more amino acids from the sequence. Substitutional 20 amino acid variants are those in which at least one residue in the sequence has been removed and a different residue inserted in its place. Typical substitutions are those made in accordance with Table 1 below.

- 14 -

TABLE 1**Suitable residues for amino acid substitutions**

	<u>Original Residue</u>	<u>Exemplary Substitutions</u>
5	Ala	Ser
	Arg	Lys
	Asn	Gln; His
	Asp	Glu
	Cys	Ser
10	Gln	Asn
	Glu	Asp
	Gly	Pro
	His	Asn; Gln
	Ile	Leu; Val
15	Leu	Ile; Val
	Lys	Arg; Gln; Glu
	Met	Leu; Ile
	Phe	Met; Leu; Tyr
	Ser	Thr
20	Thr	Ser
	Trp	Tyr
	Tyr	Trp; Phe
	Val	Ile; Leu

25 Where the F3'H is derivatised by amino acid substitution, the amino acids are generally replaced by other amino acids having like properties, such as hydrophobicity, hydrophilicity, electronegativity, bulky side chains and the like. Amino acid substitutions are typically of single residues. Amino acid insertions will usually be in the order of about 1-10 amino acid residues and deletions will range from about 1-20 residues. Preferably, deletions or insertions

30 are made in adjacent pairs, i.e. a deletion of two residues or insertion of two residues.

- 15 -

~~The amino acid variants referred to above may readily be made using peptide synthetic~~
techniques well known in the art, such as solid phase peptide synthesis (Merrifield, 1964) and
the like, or by recombinant DNA manipulations. Techniques for making substitution
mutations at predetermined sites in DNA having known or partially known sequence are well
5 known and include, for example, M13 mutagenesis. The manipulation of DNA sequence to
produce variant proteins which manifest as substitutional, insertional or deletional variants are
conveniently described, for example, in Sambrook *et al.* (1989).

Other examples of recombinant or synthetic mutants and derivatives of the F3'H of the
10 present invention include single or multiple substitutions, deletions and/or additions of any
molecule associated with the enzyme such as carbohydrates, lipids and/or proteins or
polypeptides.

The terms "analogues" and "derivatives" also extend to any chemical equivalents of the F3'H,
15 whether functional or not, and also to any amino acid derivative described above. Where the
"analogues" and "derivatives" of this aspect of the present invention are non-functional, they
may act as agonists or antagonists of F3'H activity. For convenience, reference to "F3'H"
herein includes reference to any derivatives, including parts, mutants, fragments, homologues
or analogues thereof.

20

The present invention is exemplified using nucleic acid sequences derived from petunia,
carnation, rose, snapdragon, arabidopsis, chrysanthemum, lisianthus, torenia, morning glory
and gentian, since these represent the most convenient and preferred sources of material to
date. However, one skilled in the art will immediately appreciate that similar sequences can
25 be isolated from any number of sources such as other plants or certain microorganisms.
Examples of other plants include, but are not limited to, maize, tobacco, cornflower,
pelargonium, apple, gerbera and african violet. All such nucleic acid sequences encoding
directly or indirectly a flavonoid pathway enzyme and in particular F3'H, regardless of their
source, are encompassed by the present invention.

30

- 16 -

The nucleic acid molecules contemplated herein may exist in either orientation alone or in combination with a vector molecule, for example an expression-vector. The term vector molecule is used in its broadest sense to include any intermediate vehicle for the nucleic acid molecule, capable of facilitating transfer of the nucleic acid into the plant cell and/or facilitating integration into the plant genome. An intermediate vehicle may, for example, be adapted for use in electroporation, microprojectile bombardment, *Agrobacterium*-mediated transfer or insertion via DNA or RNA viruses. The intermediate vehicle and/or the nucleic acid molecule contained therein may or may not need to be stably integrated into the plant genome. Such vector molecules may also replicate and/or express in prokaryotic cells. Preferably, the vector molecules or parts thereof are capable of integration into the plant genome. The nucleic acid molecule may additionally contain a promoter sequence capable of directing expression of the nucleic acid molecule in a plant cell. The nucleic acid molecule and promoter may also be introduced into the cell by any number of means such as those described above.

15

In accordance with the present invention, a nucleic acid molecule encoding a F3'H or a derivative or part thereof may be introduced into a plant in either orientation to allow, permit or otherwise facilitate manipulation of levels of production of mRNA in the cytoplasm and/or production of enzyme from the mRNA, thereby providing a means either to convert DHK and/or other suitable substrates, if synthesised in the plant cell, ultimately into anthocyanin derivatives of anthocyanidins such as cyanidin and/or peonidin, or alternatively to inhibit such conversion of metabolites by reducing or eliminating endogenous or existing F3'H activity. The production of mRNA in the cytoplasm and/or production of enzyme from the mRNA, is referred to herein as "expression". The production of anthocyanins contributes to the production of a red or blue flower colour. Expression of the nucleic acid molecule in either orientation in the plant may be constitutive, inducible or developmental, and may also be tissue-specific.

According to this aspect of the present invention there is provided a method for producing a transgenic plant capable of synthesizing F3'H or functional derivatives thereof, said method

30

- 17 -

comprising stably transforming a cell of a suitable plant with a nucleic acid molecule which comprises a sequence of nucleotides encoding said F3'H, under conditions permitting the eventual expression of said nucleic acid molecule, regenerating a transgenic plant from the cell and growing said transgenic plant for a time and under conditions sufficient to permit the
5 expression of the nucleic acid molecule. The transgenic plant may thereby produce elevated levels of F3'H activity relative to the amount expressed in a comparable non-transgenic plant.

Another aspect of the present invention contemplates a method for producing a transgenic plant with reduced endogenous or existing F3'H activity, said method comprising stably
10 transforming a cell of a suitable plant with a nucleic acid molecule which comprises a sequence of nucleotides encoding or complementary to a sequence encoding F3'H, regenerating a transgenic plant from the cell and where necessary growing said transgenic plant under conditions sufficient to permit the expression of the nucleic acid molecule.

15 Yet another aspect of the present invention contemplates a method for producing a genetically modified plant with reduced endogenous or existing F3'H activity, said method comprising altering the F3'H gene through modification of the endogenous sequences via homologous recombination from an appropriately altered F3'H gene or derivative or part thereof introduced into the plant cell, and regenerating the genetically modified plant from the cell.

20

In accordance with these aspects of the present invention the preferred nucleic acid molecules are substantially as set forth in SEQ ID NO:1, 3, 5, 7, 14, 16, 18, 20, 22, 24, or the coding region of 9, or have at least about 60% similarity thereto, or are capable of hybridising thereto under low stringency conditions.

25

In a preferred embodiment, the present invention contemplates a method for producing a transgenic flowering plant exhibiting altered flower colour, said method comprising stably transforming a cell of a suitable plant with a nucleic acid molecule of the present invention, regenerating a transgenic plant from the cell and growing said transgenic plant for a time and
30 under conditions sufficient to permit the expression of the nucleic acid molecule into the F3'H

- 18 -

enzyme. Alternatively, said method may comprise stably transforming a cell of a suitable plant with a nucleic acid molecule of the present invention or its complementary sequence, regenerating a transgenic plant from the cell and growing said transgenic plant for a time and under conditions sufficient to alter the level of activity of the endogenous or existing F3'H. 5 Preferably, the altered level would be less than the endogenous or existing level of F3'H activity in a comparable non-transgenic plant.

In a related embodiment, the present invention contemplates a method for producing a flowering plant exhibiting altered flower colour, said method comprising alteration of the 10 F3'H gene through modification of the endogenous sequences *via* homologous recombination from an appropriately altered F3'H gene or derivative thereof introduced into the plant cell and regenerating the genetically modified plant from the cell.

The nucleic acid molecules of the present invention may or may not be developmentally 15 regulated. Preferably, the modulation of levels of 3'-hydroxylated anthocyanins leads to altered flower colour which includes the production of red flowers or other colour shades depending on the physiological conditions of the recipient plant. By "recipient plant" is meant a plant capable of producing a substrate for the F3'H enzyme, or producing the F3'H enzyme itself, and possessing the appropriate physiological properties and genotype required 20 for the development of the colour desired. This may include but is not limited to petunia, carnation, chrysanthemum, rose, snapdragon, tobacco, cornflower, pelargonium, lisianthus, gerbera, apple, iris, lily, african violet, gentian, torenia and Japanese morning glory.

Accordingly, the present invention extends to a method for producing a transgenic plant 25 capable of modulating levels of 3'-hydroxylated anthocyanins, said method comprising stably transforming a cell or group of cells of a suitable plant with a nucleic acid molecule comprising a sequence of nucleotides encoding, or complementary to a sequence encoding, F3'H or a derivative thereof, and regenerating a transgenic plant from said cell or cells.

30 One skilled in the art will immediately recognise the variations applicable to the methods of

~~the present invention, such as increasing or decreasing the level of enzyme activity of the~~
enzyme naturally present in a target plant leading to differing shades of colours.

The present invention, therefore, extends to all transgenic plants containing all or part of the
5 nucleic acid module of the present invention and/or any homologues or related forms thereof
or antisense forms of any of these and in particular those transgenic plants which exhibit
altered flower colour. The transgenic plants may contain an introduced nucleic acid molecule
comprising a nucleotide sequence encoding or complementary to a sequence encoding F3'H.
Generally, the nucleic acid would be stably introduced into the plant genome, although the
10 present invention also extends to the introduction of the F3'H nucleotide sequence within an
autonomously-replicating nucleic acid sequence such as a DNA or RNA virus capable of
replicating within the plant cell. The invention also extends to seeds from such transgenic
plants. Such seeds, especially if coloured, will be useful as proprietary tags for plants.

15 A further aspect of the present invention is directed to recombinant forms of F3'H. The
recombinant forms of the enzymes will provide a source of material for research to develop,
for example, more active enzymes and may be useful in developing *in vitro* systems for
production of coloured compounds.

20 Still a further aspect of the present invention contemplates the use of the genetic sequences
described herein in the manufacture of a genetic construct capable of use in modulating levels
of 3'-hydroxylated anthocyanins in a plant or cells of a plant.

Yet a further aspect of the present invention provides flowers and in particular cut flowers,
25 from the transgenic plants herein described, exhibiting altered flower colour.

Another aspect of the present invention is directed to a nucleic acid molecule comprising a
sequence of nucleotides encoding or complementary to a sequence encoding, a F3'H or a
derivative thereof wherein said nucleic acid molecule is capable of being expressed in a plant
30 cell. The term "expressed" is equivalent to the term "expression" as defined above.

The nucleic acid molecules according to this and other aspects of the invention allow, permit or otherwise facilitate increased efficiency in modulation of 3'-hydroxylated anthocyanins relative to the efficiency of the pCGP619 cDNA insert contained in plasmid pCGP809, disclosed in International Patent Application No. PCT/AU93/00127 [WO 93/20206]. The
 5 term "plant cell" includes a single plant cell or a group of plant cells such as in a callus, plantlet or plant or parts thereof including flowers and seeds.

Another aspect of the present invention provides a nucleic acid molecule comprising a sequence of nucleotides encoding or complementary to a sequence of nucleotides encoding
 10 a F3'H, wherein the translation of the said nucleic acid molecule comprises the amino acid sequence RPPNSGA. Preferably, the translation of the said nucleic acid molecule comprises the amino acid sequence RPPNSGAXHXAYNYXDL and still more preferably the translation of the said nucleic acid molecule comprises the amino acid sequence RPPNSGAXHXAYNYXDL[X]_nGGEK, where X represents any amino acid and [X]_n
 15 represents an amino acid sequence of from 0 to 500 amino acids.

The present invention is further described by reference to the following non-limiting Figures and Examples.

20 In the Figures:

Figures 1a and 1b are schematic representations of the flavonoid biosynthesis pathways in *P. hybrida* flowers showing the enzymes and genetic loci involved in the conversions. Enzymes involved in the pathway have been indicated as follows: PAL = phenylalanine
 25 ammonia-lyase; C4H = cinnamate 4-hydroxylase; 4CL = 4-coumarate: CoA ligase; CHS = chalcone synthase; CHI = chalcone isomerase; F3H = flavanone 3-hydroxylase; F3'H = flavonoid 3'-hydroxylase; F3'5'H = flavonoid 3'5' hydroxylase; FLS = flavonol synthase; DFR = dihydroflavonol-4-reductase; ANS = anthocyanin synthase; 3GT = UDP-glucose: anthocyanin-3-glucoside; 3RT = UDP-rhamnose: anthocyanidin-3-glucoside
 30 rhamnosyltransferase; ACT = anthocyanidin-3-rutinoside acyltransferase; 5GT = UDP-

- 21 -

glucose: anthocyanin-5-glucosyltransferase; 3'-OMT=anthocyanin-O-methyltransferase; 3',
 5' OMT=anthocyanin 3', 5' O-methyltransferase. Three flavonoids in the pathway are
 indicated as: P-3-G= pelargonidin-3-glucoside; DHM=dihydomyricetin;
 DHQ=dihydroquercetin. The flavonol, myricetin, is only produced at low levels and the
 5 anthocyanin, pelargonidin, is rarely produced in *P. hybrida*.

- Figure 2** is a diagrammatic representation of the plasmid pCGP161 containing a cDNA clone (F1) representing the cinnamate-4-hydroxylase from *P. hybrida*. ³²P-labelled fragments of the 0.7 kb EcoRI/XhoI fragment were used to probe the Old Glory Red petal cDNA library.
- 10 For details, refer to Example 4. Abbreviations are as follows: Amp = the ampicillin resistance gene; ori = origin of replication; T3 = recognition sequence for T3 RNA polymerase; T7 = recognition sequence for T7 RNA polymerase. Restriction enzyme sites are also marked.
- 15 **Figure 3** is a diagrammatic representation of the plasmid pCGP602 containing a cDNA clone (617) representing a flavonoid 3'5' hydroxylase (Hf1) from *P. hybrida*. ³²P-labelled fragments of the 1.6 kb BspHI/EspI fragment containing the Hf1 coding region were used to probe the Old Glory Red petal cDNA library. For details, refer to Example 4. Abbreviations are as follows: Amp = the ampicillin resistance gene; ori = origin of replication; T3 =
- 20 recognition sequence for T3 RNA polymerase; T7 = recognition sequence for T7 RNA polymerase. Restriction enzyme sites are also marked.

Figure 4 is a diagrammatic representation of the plasmid pCGP175 containing a cDNA clone (H2) representing a flavonoid 3'5' hydroxylase (Hf2) from *P. hybrida*. ³²P-labelled

25 fragments of the 1.3 kb EcoRI/XhoI and 0.5 kb XhoI fragments which together contain the Hf2 coding region were used to probe the Old Glory Red petal cDNA library. For details, refer to Example 4. Abbreviations are as follows: Amp = the ampicillin resistance gene; ori = origin of replication; T3 = recognition sequence for T3 RNA polymerase; T7 = recognition sequence for T7 RNA polymerase. Restriction enzyme sites are also marked.

Figure 5 is a diagrammatic representation of the plasmid pCGP619 containing the 651 cDNA clone representing a cytochrome P450 from *P. hybrida*. ³²P-labelled fragments of the 1.8 kb EcoRI/XhoI fragment were used to probe the Old Glory Red petal cDNA library. For details, refer to Example 4. Abbreviations are as follows: Amp = the ampicillin resistance gene; ori = origin of replication; T3 = recognition sequence for T3 RNA polymerase; T7 = recognition sequence for T7 RNA polymerase. Restriction enzyme sites are also marked.

Figure 6 is a representation of an autoradiograph of an RNA blot probed with ³²P-labelled fragments of the OGR-38 cDNA clone contained in pCGP1805 (see Example 6). Each lane contained a 20 µg sample of total RNA isolated from the flowers or leaves of plants of a V23 (*ht1/ht1*) x VR (*Ht1/ht1*) backcross population. A 1.8 kb transcript was detected in the VR-like (*Ht1/ht1*) flowers that contained high levels of quercetin (Q+)(lanes 9 - 14). The same size transcript was detected at much lower levels in the V23-like (*ht1/ht1*) flowers that contained little or no quercetin (Q-) (lanes 3-8). A reduced level of transcript was also detected in VR leaves (lane 1) and V23 petals (lane 2). This is described in Example 5.

Figure 7 is a diagrammatic representation of the yeast expression plasmid pCGP1646 (see Example 7). The OGR-38 cDNA insert from pCGP1805 was cloned in a "sense" orientation behind the yeast glyceraldehyde-3-phosphate dehydrogenase promoter (PGAP) in the expression vector pYE22m. TRP1 = Trp1 gene, IR1 = inverted repeat of 2 µm plasmid, TGAP = terminator sequence from the yeast glyceraldehyde-3-phosphate dehydrogenase gene. Restriction enzyme sites are also marked.

Figure 8 is a diagrammatic representation of the binary plasmid pCGP1867 (described in Example 8). The *Ht1* cDNA insert (OGR-38) from pCGP1805 was cloned in a "sense" orientation behind the Mac promoter in the expression vector of pCGP293. Abbreviations are as follows: LB = left border; RB = right border; Gm = the gentamycin resistance gene; 35S = the promoter region from the Cauliflower Mosaic Virus 35S gene; nptII = the neomycin phosphotransferase II gene; tml3' = the terminator region from the *tml* gene of *Agrobacterium*; mas3' = the terminator region from the mannopine synthase gene of

- 23 -

Agrobacterium; ori pRi = a broad host range origin of replication from an *Agrobacterium rhizogenes* plasmid; oriColE1 = a high copy origin of replication from a Colcinin E1 plasmid. Restriction enzyme sites are also marked.

5 **Figure 9** is a diagrammatic representation of the binary plasmid pCGP1810, preparation of which is described in Example 13. The KC-1 cDNA insert from pCGP1807 (see Example 12) was cloned in a "sense" orientation behind the Mac promoter in the expression vector of pCGP293. Abbreviations are as follows: LB = left border; RB = right border; Gm = the gentamycin resistance gene; 35S = the promoter region from the Cauliflower Mosaic Virus
10 35S gene; nptII = the neomycin phosphotransferase II gene; tml3' = the terminator region from the *tml* gene of *Agrobacterium*; mas3' = the terminator region from the mannopine synthase gene of *Agrobacterium*; ori pRi = a broad host range origin of replication from a plasmid from *Agrobacterium rhizogenes*; oriColE1 = a high copy origin of replication from a Colcinin E1 plasmid. Restriction enzyme sites are also marked.

15

Figure 10 is a diagrammatic representation of the binary plasmid pCGP1813, construction of which is described in Example 14. The KC-1 cDNA insert from pCGP1807 (see Example 12) was cloned in a "sense" orientation between the mac promoter and mas terminator. The Mac: KC-1: mas expression cassette was subsequently cloned into the binary vector
20 pWTT2132. Abbreviations are as follows: Tet= the tetracycline resistance gene; LB= left border; RB=right border, *surB*=the coding region and terminator sequence from the acetolactate synthase gene; 35S= the promoter region from the cauliflower mosaic virus 35S gene, mas3'=the terminator region from the mannopine synthase gene of *Agrobacterium*; pVS1 = a broad host range origin of replication from a plasmid from *Pseudomonas*
25 *aeruginosa*, pACYCori= modified replicon from pACYC184 from *E. coli*. Restriction enzyme sites are also marked.

Figure 11 is a representation of an autoradiograph of a Southern blot probed with ³²P-labelled fragments of the Am3Ga differential display PCR fragment (as described in Example
30 16). Each lane contained a 10 µg sample of *EcoRV*-digested genomic DNA isolated from N8

- 24 -

(Eos⁺), K16 (eos⁻) or plants of an K16 x N8 F₂ population. Hybridizing bands were detected in the genomic DNA from cyanidin-producing plants (indicated with "+") (Lanes 1, 3, 4, 5, 6, 7, 9, 10, 12 and 15). No specific hybridization was observed in the genomic DNA samples from non-cyanidin-producing plants (indicated with "-") (Lanes 2, 8, 11, 13 and 14).

Figure 12 is a representation of an autoradiograph of an RNA blot probed with ³²P-labelled fragments of the Am3Ga differential display PCR fragment. Each lane contained a 10 µg sample of total RNA isolated from the flowers or leaves of plants of an N8 (Eos⁺) x K16 (eos⁻) F₂ population. A 1.8 kb transcript was detected in the K16 x N8 F₂ flowers that produced cyanidin (cyanidin +) (plants #1, #3, #4, #5 and #8). No transcript was detected in the K16 x N8 F₂ flowers that did not produce cyanidin (cyanidin -) (plants #6, #11, #12) or in a leaf sample (#13L) from an K16 x N8 F₂ plant that produced cyanidin in the flowers. Details are provided in Example 17.

15

Figure 13 is a diagrammatic representation of the binary plasmid pCGP250, construction of which is described in Example 20. The sdf3'H cDNA insert, containing the nucleotides 1 through to 1711 (SEQ ID NO:5) from pCGP246 (see Example 18), was cloned in a "sense" orientation behind the Mac promoter in the expression vector of pCGP293. Abbreviations are as follows: LB = left border; RB = right border; Gm = the gentamycin resistance gene; 35S = the promoter region from the Cauliflower Mosaic Virus 35S gene; nptII = the neomycin phosphotransferase II gene; tml3' = the terminator region from the tml gene of *Agrobacterium*; mas3' = the terminator region from the mannopine synthase gene of *Agrobacterium*; ori pRi = a broad host range origin of replication from a plasmid from *Agrobacterium rhizogenes*; oriColE1 = a high copy origin of replication from a Colicin E1 plasmid. Restriction enzyme sites are also marked.

Figure 14 is a diagrammatic representation of the binary plasmid pCGP231, construction of which is described in Example 20. The sdf3'H cDNA insert, containing the nucleotides 104 through to 1711 (SEQ ID NO:5) from pCGP246, was cloned in a "sense" orientation behind

30

- 25 -

- the Mac promoter in the expression vector of pCGP293. Abbreviations are as follows: LB = left border; RB = right border; Gm = the gentamycin resistance gene; 35S = the promoter region from the Cauliflower Mosaic Virus 35S gene; nptII = the neomycin phosphotransferase II gene; tml3' = the terminator region from the tml gene of *Agrobacterium*; mas3' = the terminator region from the mannopine synthase gene of *Agrobacterium*; ori pRi = a broad host range origin of replication from a plasmid from *Agrobacterium rhizogenes*; oriColE1 = a high copy origin of replication from a Colcinin E1 plasmid. Restriction enzyme sites are also marked.
- 10 **Figure 15** is a diagrammatic representation of the binary plasmid pBI-Tt7-2. The 6.5 kb EcoRI/SalI Tt7 genomic fragment from E-5 was cloned into EcoRI/SalI-cut pBI101, replacing the resident GUS gene. The orientation of the Tt7 (F3'H) gene as indicated (5' to 3') was determined through DNA sequencing. Abbreviations are as follows: LB = left border; RB = right border; nos 5' = the promoter region from the nopaline synthase gene of *Agrobacterium*; nptII = the coding region of the neomycin phosphotransferase II gene; nos 3' = the terminator region from the nopaline synthase gene of *Agrobacterium*; nptI = the coding region of the neomycin phosphotransferase I gene. Restriction enzyme sites are also marked.
- 20 **Figure 16** is a diagrammatic representation of the binary plasmid pCGP2166, construction of which is described in Example 26. The rose #34 cDNA insert from pCGP2158 (see Example 25) was cloned in a "sense" orientation behind the Mac promoter in the expression vector of pCGP293. Abbreviations are as follows: LB = left border; RB = right border; Gm = the gentamycin resistance gene; 35S = the promoter region from the Cauliflower Mosaic Virus 35S gene; nptII = the neomycin phosphotransferase II gene; tml3' = the terminator region from the tml gene of *Agrobacterium*; mas3' = the terminator region from the mannopine synthase gene of *Agrobacterium*; ori pRi = a broad host range origin of replication from a plasmid from *Agrobacterium rhizogenes*; oriColE1 = a high copy origin of replication from a Colcinin E1 plasmid. Restriction enzyme sites are also marked.

- 26 -

Figure 17 is a diagrammatic representation of the binary plasmid pCGP2169 construction of which is described in Example 27. The rose #34 cDNA insert from pCGP2158 was cloned in a "sense" orientation between the CaMV35S promoter and the ocs terminator. The 35S: rose #34: ocs expression cassette was subsequently cloned into the binary vector pWTT2132.

5 Abbreviations are as follows: Tet= the tetracycline resistance gene; LB= left border; RB=right border; surB=the coding region and terminator sequence from the acetolactate synthase gene; 35S=the promoter region from the cauliflowe mosaic virus 35S gene, ocs=terminator region from the octopine synthase gene from *Agrobacterium*; pVS1=a broad host range origin of replication from a plasmid from *Pseudomonas aeruginosa*,

10 pACYCori=modified replicon from pACYC184 from *E. coli*. Restriction enzyme sites are also marked.

Figure 18 is a diagrammatic representation of the binary plasmid pLN85, construction of which is described in Example 28. The chrysanthemum RM6i cDNA insert from pCHRM1

15 was cloned in "anti-sense" orientation behind the promoter from the Cauliflower Mosaic Virus 35S gene (35S). Other abbreviations are as follows: LB = left border; RB = right border; ocs3' = the terminator region from the octopine synthase gene of *Agrobacterium*; pnos:nptII:nos 3' = the expression cassette containing the promoter region from the nopaline synthase gene of *Agrobacterium*; the coding region of the neomycin phosphotransferase II

20 gene and the terminator region from the nopaline synthase gene of *Agrobacterium*; oriT = origin of transfer of replication; trfA* = a trans-acting replication function; oriColE1 = a high copy origin of replication from a Colicin E1 plasmid; Tn7SpR/StR = the spectinomycin and streptomycin resistance genes from transposon Tn7; oriVRK2 = a broad host range origin of replication from plasmid RK2. Restriction enzyme sites are also marked.

25

Figure 19 is a diagrammatic representation of the yeast expression plasmid pYTHT6, construction of which is described in Example 30. The THT6 cDNA insert from pTHT6 was cloned in a "sense" orientation behind the yeast glyceraldehyde-3-phosphate dehydrogenase promoter (PGAP) in the expression vector pYE22m. Abbreviations are as follows: TRP1 =

30 Trp1 gene; IR1 = inverted repeat of 2 μ m plasmid; TGAP = the terminator sequence from

- 27 -

~~the yeast glyceraldehyde-3-phosphate dehydrogenase gene. Restriction-enzyme sites are also~~
marked.

5 Amino acid abbreviations used throughout the specification are shown in Table 2, below.

- 28 -

TABLE 2**Amino acid abbreviations**

	Amino acid	3-letter	single-letter
5	////////////////////		
	L-alanine	Ala	A
	L-arginine	Arg	R
	L-asparagine	Asn	N
	L-aspartic acid	Asp	D
10	L-cysteine	Cys	C
	L-glutamine	Gln	Q
	L-glutamic acid	Glu	E
	L-glycine	Gly	G
	L-histidine	His	H
15	L-isoleucine	Ile	I
	L-leucine	Leu	L
	L-lysine	Lys	K
	L-methionine	Met	M
	L-phenylalanine	Phe	F
20	L-proline	Pro	P
	L-serine	Ser	S
	L-threonine	Thr	T
	L-tryptophan	Trp	W
	L-tyrosine	Tyr	Y
25	L-valine	Val	V
	////////////////////		

Table 3 provides a summary of the SEQ ID NO's assigned to the sequences referred to herein:

TABLE 3

5	Sequence	Species	SEQ ID NO
	cDNA insert of pCGP1805		
	corresponding amino acid sequence	Petunia	SEQ ID NO:1
	cDNA insert of pCGP1807		
10	corresponding amino acid sequence	Petunia	SEQ ID NO:2
	cDNA insert of pCGP1807		
	corresponding amino acid sequence	Carnation	SEQ ID NO:3
	cDNA insert of pCGP246		
	corresponding amino acid sequence	Carnation	SEQ ID NO:4
	cDNA insert of pCGP246		
	corresponding amino acid sequence	Snapdragon	SEQ ID NO:5
	cDNA partial sequence		
	corresponding amino acid sequence	Snapdragon	SEQ ID NO:6
	cDNA partial sequence		
15	corresponding amino acid sequence	Arabidopsis	SEQ ID NO:7
	genomic sequence		
	corresponding amino acid sequence	Arabidopsis	SEQ ID NO:8
	genomic sequence		
	corresponding amino acid sequence	Arabidopsis	SEQ ID NO:9
	for exon I		
	corresponding amino acid sequence	Arabidopsis	SEQ ID NO:10
	for exon II		
20	corresponding amino acid sequence	Arabidopsis	SEQ ID NO:11
	for exon III		
	corresponding amino acid sequence	Arabidopsis	SEQ ID NO:12
	for exon IV		
	corresponding amino acid sequence	Arabidopsis	SEQ ID NO:13
25	cDNA insert of pCGP2158	Rose	SEQ ID NO:14
	corresponding amino acid sequence	Rose	SEQ ID NO:15
	cDNA insert of pCHRM1		
	corresponding amino acid sequence	Chrysanthemum	SEQ ID NO:16
	THT cDNA sequence		
	corresponding amino acid sequence	Chrysanthemum	SEQ ID NO:17
	THT cDNA sequence		
	corresponding amino acid sequence	Torenia	SEQ ID NO:18
30	corresponding amino acid sequence	Torenia	SEQ ID NO:19

- 30 -

	MHT 85 cDNA sequence	Jap. Morning Glory	SEQ ID NO:20
	corresponding amino acid sequence	Jap. Morning Glory	SEQ ID NO:21
	GHT13 cDNA sequence	Gentian	SEQ ID NO:22
	corresponding amino acid sequence	Gentian	SEQ ID NO:23
5	cDNA insert of pL3-6	Lisianthus	SEQ ID NO:24
	corresponding amino acid sequence	Lisianthus	SEQ ID NO:25
	cDNA sequence from WO 93/20206	Petunia	SEQ ID NO:26
	oligonucleotide polyT-anchA		SEQ ID NO:27
	oligonucleotide polyT-anchC		SEQ ID NO:28
10	oligonucleotide polyT-anchG		SEQ ID NO:29
	conserved amino acid primer region		SEQ ID NO:30
	corresponding oligonucleotide sequence		SEQ ID NO:31
	conserved amino acid primer region		SEQ ID NO:32
	corresponding oligonucleotide sequence		SEQ ID NO:33
15	oligonucleotide primer Pet Haem-New		SEQ ID NO:34
	conserved amino acid primer region		SEQ ID NO:35
	corresponding oligonucleotide sequence		SEQ ID NO:36
	oligonucleotide Snapred Race A		SEQ ID NO:37
	oligonucleotide Snapred Race C		SEQ ID NO:38
20	oligonucleotide poly-C Race		SEQ ID NO:39
	oligonucleotide primer Pet Haem		SEQ ID NO:40

////////////////////////////////////

25 The disarmed microorganism *Agrobacterium tumefaciens* strain AGL0 separately containing the plasmids pCGP1867, pCGP1810 and pCGP231 were deposited with the Australian Government Analytical Laboratories, 1 Suakin Street, Pymble, New South Wales, 2037, Australia on 23 February, 1996 and were given Accession Numbers 96/10967, 96/10968 and 96/10969, respectively.

ISOLATION OF FLAVONOID 3'-HYDROXYLASE AND RELATED NUCLEIC ACID SEQUENCES

5 EXAMPLE 1-Plant Material

Petunia

The *Petunia hybrida* varieties used are presented in Table 4.

TABLE 4

10

Plant variety	Properties	Source/Reference
Old Glory Blue (OGB)	F ₁ Hybrid	Ball Seed, USA
Old Glory Red (OGR)	F ₁ Hybrid	Ball Seed, USA
15 V23	<i>An1, An2, An3, An4, An6, An8, An9, An10, ph1, Hf1, Hf2, ht1, Rt, po, Bl, Fl</i>	Wallroth <i>et al.</i> (1986) Doodeman <i>et al.</i> (1984)
R51	<i>An1, An2, An3, an4, An6, An8, An9, An10, An11, Ph1, hf1, hf2, Ht1, rt, Po, bl, fl</i>	Wallroth <i>et al.</i> (1986) Doodeman <i>et al.</i> (1984)
VR	V23 x R51 F ₁ Hybrid	
SW63	<i>An1, An2, An3, an4, An6, An8, An9, An10, An11, Ph1, Ph2, Ph5, hf1, hf2, ht1, ht2, po, mfl, fl</i>	I.N.R.A., Dijon, Cedex France
Skr4	<i>An1, An2, An3, An4, An6, An11, hf1, hf2, ht1, Ph1, Ph2, Ph5, rt, Po, Mfl, Mf2, fl</i>	I.N.R.A., Dijon, Cedex France
20 Skr4 x SW63	F ₁ Hybrid	

Plants were grown in specialised growth rooms with a 14 hour day length at a light intensity of 10,000 lux and a temperature of 22°C to 26°C.

25 Carnation

Flowers of *Dianthus caryophyllus* cv. Kortina Chanel were obtained from Van Wyk and Son

Flower Supply, Victoria.

Dianthus caryophyllus flowers were harvested at developmental stages defined as follows:

- 5 Stage 1: Closed bud, petals not visible.
- Stage 2: Flower buds opening: tips of petals visible.
- Stage 3: Tips of nearly all petals exposed. "Paint-brush stage".
- Stage 4: Outer petals at 45° angle to stem.
- Stage 5: Flower fully open.

10

Snapdragon

The *Antirrhinum majus* lines used were derived from the parental lines K16 (*eos*⁻) and N8 (*Eos*⁺). A strict correlation exists between F3'H activity and the *Eos* gene which is known to control the 3'-hydroxylation of flavones, flavonols and anthocyanins (Forkmann and Stotz, 1981). K16 is a homozygous recessive mutant lacking F3'H activity while N8 is wild type for F3'H activity. These lines are similar, though not isogenic. Both parental lines and the seed from a selfed (K16 x N8) F₁ plant were obtained from Dr C. Martin (John Innes Centre, Norwich, UK).

20 **Arabidopsis**

The *Arabidopsis thaliana* lines Columbia (*Tt7*), Landsberg erecta (*Tt7*) and NW88 (*tt7*) were obtained from the Nottingham Arabidopsis Stock Centre. Wild-type *A. thaliana* (*Tt7*) seeds have a characteristic brown colour. Seeds of *tt7* mutants have pale brown seeds and the plants are characterized by a reduced anthocyanin content in leaves (Koornneef et al., 1982). 25 *Tt7* plants produce cyanidin while *tt7* mutants accumulate pelargonidin, indicating that the *Tt7* gene controls flavonoid 3'-hydroxylation.

Rose

Flowers of *Rosa hybrida* cv. Kardinal were obtained from Van Wyk and Son Flower Supply, 30 Victoria.

- 33 -

Stages of *Rosa hybrida* flower development were defined as follows:

- Stage 1: Unpigmented, tightly closed bud (10-12 mm high; 5 mm wide).
- Stage 2: Pigmented, tightly closed bud (15 mm high ; 9 mm wide).
- 5 Stage 3: Pigmented, closed bud; sepals just beginning to open (20-25 mm high; 13-15 mm wide)
- Stage 4: Flower bud beginning to open; petals heavily pigmented; sepals have separated (bud is 25-30 mm high and 18 mm wide).
- Stage 5: Sepals completely unfolded; some curling. Petals are heavily pigmented and
10 unfolding (bud is 30-33 mm high and 20 mm wide).

Chrysanthemum

Stages of *Chrysanthemum* flower development were defined as follows:

- 15 Stage 0: No visible flower bud.
- Stage 1: Flower bud visible: florets completely covered by the bracts.
- Stage 2: Flower buds opening: tips of florets visible.
- Stage 3: Florets tightly overlapped.
- Stage 4: Tips of nearly all florets exposed; outer florets opening but none horizontal.
- 20 Stage 5: Outer florets horizontal.
- Stage 6: Flower approaching maturity.

EXAMPLE 2-Bacterial Strains

25 The *Escherichia coli* strains used were:

DH5 α *supE44*, Δ (*lacZ*YA-ArgF)U169, ϕ 80*lacZ* Δ M15, *hsdR17* (*r_k*⁻, *m_k*⁺), *recA1*,
endA1, *gyrA96*, *thi-1*, *relA1*, *deoR* (Hanahan, 1983 and BRL, 1986).

30 XL1-Blue MRF' Δ (*mcrA*)183, Δ (*mcrCB*-*hsdSMR*-*mrr*)173, *endA1*, *supE44*, *thi-1*,

- 34 -

- ~~recA1, gyrA96, relA1, lac[F' proAB, lacIqZΔM15, Tn10(Tet^r)]^c~~
(Stratagene)
- XL1-Blue ~~supE44, hsdR17 (rk⁻, mk⁺), recA1, endA1, gyrA96, thi-1, relA1,~~
5 ~~lac[F' proAB, lacIq, lacZΔM15, Tn10(tet^r)]~~
- SOLR e14⁻ (~~mcrA~~), Δ(~~mcrCB-hsdSMR-mrr~~)171, ~~sbcC, recB, recJ,~~
10 ~~umuC::Tn5(kan^r), uvrC, lac, gyrA96, thi-1, relA1, [F' proAB,~~
~~lacIqZΔM15], Su⁻ (non-suppressing) (Stratagene)~~
- DH10 B(Zip) F⁻~~mcrA~~, Δ(~~mrr-hsdRMS-mcrBC~~), ø80d ~~lacZΔM15, ΔlacX74,~~
~~deoR, recA1, araD139, Δ(ara, leu)7697, galU, galKl^λ, rspL,~~
~~nupG~~
- 15 Y1090r- ΔlacU169, (Δlon)?, ~~araD139, strA, supF, mcrA,~~
~~trpC22::Tn10(Tet^r) [pMC9 Amp^r, Tet^r], mcrB, hsdR~~

The disarmed *Agrobacterium tumefaciens* strain AGL0 (Lazo *et al.*, 1991) was obtained from R. Ludwig (Department of Biology, University of California, Santa Cruz, USA).

20

The cloning vector pBluescript was obtained from Stratagene.

Transformation of the *E. coli* strain DH5α cells was performed according to the method of Inoue *et al.* (1990).

25

EXAMPLE 3-General methods

³²P-Labeling of DNA Probes

DNA fragments (50 to 100 ng) were radioactively labelled with 50 μCi of [α-³²P]-dCTP
30 using an oligolabelling kit (Bresatec). Unincorporated [α-³²P]-dCTP was removed by chromatography on a Sephadex G-50 (Fine) column.

DNA Sequence Analysis

DNA sequencing was performed using the PRISM™Ready Reaction Dye Primer Cycle Sequencing Kits from Applied Biosystems. The protocols supplied by the manufacturer were followed. The cycle sequencing reactions were performed using a Perkin Elmer PCR machine (GeneAmp PCR System9600) and run on an automated 373A DNA sequencer (Applied Biosystems).

Homology searches against Genbank, SWISS-PROT and EMBL databases were performed using the FASTA and TFASTA programs (Pearson and Lipman, 1988) or BLAST programs (Altschul *et al.*, 1990). Percentage sequence similarities were obtained using the LFASTA program (Pearson and Lipman, 1988). In all cases ktup values of 6 for nucleotide sequence comparisons and 2 for amino acid sequence comparisons were used, unless otherwise specified.

Multiple sequence alignments (ktup value of 2) were performed using the ClustalW program incorporated into the MacVector™6.0 application (Oxford Molecular Ltd.).

EXAMPLE 4- Isolation of a flavonoid 3'-hydroxylase (F3'H) cDNA clone corresponding to the Ht1 locus from *P. hybrida* cv. Old Glory Red

In order to isolate a cDNA clone that was linked to the Ht1 locus and that represented the flavonoid 3'-hydroxylase (F3'H) in the petunia flavonoid pathway, a petal cDNA library was prepared from RNA isolated from stages 1 to 3 of Old Glory Red (OGR) petunia flowers. OGR flowers contain cyanidin based pigments and have high levels of flavonoid 3'-hydroxylase activity. The OGR cDNA library was screened with a mixture of ³²P-labelled fragments isolated from three cytochrome P450 cDNA clones known to be involved in the flavonoid pathway and from one cytochrome P450 cDNA clone (651) that had flavonoid 3'-hydroxylase activity in yeast. These included a petunia cDNA clone representing the cinnamate-4-hydroxylase (C4H) and two petunia cDNA clones (coded by the Hf1 and Hf2 loci) representing flavonoid 3' 5'-hydroxylase (F3' 5'H) (Holton *et al.*, 1993).

Construction of Petunia cv. OGR cDNA library

Total RNA was isolated from the petal tissue of *P. hybrida* cv OGR stage 1 to 3 flowers using the method of Turpen and Griffith (1986). Poly(A)⁺ RNA was selected from the total
5 RNA, using oligotex-dT[™] (Qiagen).

A ZAP-cDNA Gigapack III Gold Cloning kit (Stratagene) was used to construct a directional petal cDNA library in λ ZAP using 5 μ g of poly(A)⁺ RNA isolated from stages 1 to 3 of OGR as template. The total number of recombinants obtained was 2.46×10^6 .

10

After transfecting XL1-Blue MRF' cells, the packaged cDNA mixture was plated at 50,000 pfu per 15 cm diameter plate. The plates were incubated at 37°C for 8 hours, and the phage were eluted in 100 mM NaCl, 8 mM MgSO₄, 50 mM Tris-HCl pH 8.0, 0.01% (w/v) gelatin (Phage Storage Buffer (PSB)) (Sambrook *et al.*, 1989). Chloroform was added and the
15 phage stored at 4°C as an amplified library.

100,000 pfu of the amplified library were plated onto NZY plates (Sambrook *et al.*, 1989) at a density of 10,000 pfu per 15 cm plate after transfecting XL1-Blue MRF' cells, and incubated at 37°C for 8 hours. After incubation at 4°C overnight, duplicate lifts were taken
20 onto Colony/Plaque Screen[™] filters (DuPont) and treated as recommended by the manufacturer.

Isolation of probes

F3'5'H probes

25 The two flavonoid 3', 5' hydroxylases corresponding to the *Hf1* or *Hf2* loci isolated as described in Holton *et al.* (1993) and US Patent Number 5,349,125, were used in the screening process.

C4H cDNA clones from petunia

30 A number of cytochrome P450 cDNA clones were isolated in the screening process used to

- 37 -

isolate the two flavonoid 3', 5' hydroxylase cDNA clones corresponding to the Hf1 or Hf2 loci (Holton *et al.*, 1993; US Patent Number 5,349,125). One of these cDNA clones (F1) (contained in pCGP161) (Figure 2) was identified as representing a cinnamate 4-hydroxylase (C4H), based on sequence identity with a previously-characterised C4H clone from mung bean (Mizutani *et al.*, 1993). Sequence data was generated from 295 nucleotides at the 5' end of the petunia F1 cDNA clone. There was 83.1 % similarity with the mung bean C4H clone over the 295 nucleotides sequenced and 93.9% similarity over the predicted amino acid sequence.

10 651 cDNA clone

The isolation and identification of the 651 cDNA clone contained in pCGP619 (Figure 5) was described in the International Patent Application, having publication number W093/20206. A protein extract of yeast containing the 651 cDNA clone under the control of the yeast glyceraldehyde-3-phosphate dehydrogenase promoter of pYE22m (Tanaka *et al.*, 15 1988) exhibited F3'H activity.

Screening of OGR Library

Prior to hybridization, the duplicate plaque lifts were washed in prewashing solution (50 mM Tris-HCl pH7.5, 1 M NaCl, 1 mM EDTA, 0.1 % (w/v) sarcosine) at 65°C for 30 minutes; 20 stripped in 0.4 M sodium hydroxide at 65°C for 30 minutes; then washed in a solution of 0.2 M Tris-HCl pH 8.0, 0.1 x SSC, 0.1 % (w/v) SDS at 65°C for 30 minutes and finally rinsed in 2 x SSC, 1.0 % (w/v) SDS.

The lifts from the OGR cDNA library were screened with ³²P-labelled fragments of (1) a 25 0.7 kb EcoRI/XhoI fragment from pCGP161 containing the C4H cDNA clone (Figure 2), (2) a 1.6 kb BspHI/EspI fragment from pCGP602 containing the Hf1 cDNA clone (Figure 3), (3) a 1.3 kb EcoRI/XhoI fragment and a 0.5 kb XhoI fragment from pCGP175 containing the coding region of the Hf2 cDNA clone (Figure 4) and (4) a 1.8 kb EcoRI/XhoI fragment pCGP619 containing the 651 cDNA clone (Figure 5).

Hybridization conditions included a prehybridization step in 10% (v/v) formamide, 1 M NaCl, 10% (w/v) dextran sulphate, 1% (w/v) SDS at 42°C for at least 1 hour. The ³²P-labelled fragments (each at 1x10⁶cpm/mL) were then added to the hybridization solution and hybridization was continued at 42°C for a further 16 hours. The filters were then washed in 2 x SSC, 1% (w/v) SDS at 42°C for 2 x 1 hour and exposed to Kodak XAR film with an intensifying screen at -70°C for 16 hours.

Two hundred and thirty strongly hybridizing plaques were picked into PSB. Of these, 39 were rescreened to isolate purified plaques, using the hybridization conditions as described for the initial screening of the cDNA library. The plasmids contained in the λZAP bacteriophage vector were rescued and sequence data was generated from the 3' and 5' ends of the cDNA inserts. Based on sequence homology, 27 of the 39 were identical to the petunia cinnamate 4-hydroxylase cDNA clone, 2 of the 39 were identical to the Ht1 cDNA clone and 7 of the 39 did not represent cytochrome P450s. The remaining 3 cDNA clones (designated as OGR-27, OGR-38, OGR-39) represented "new" cytochrome P450s, compared to the cytochrome P450 clones used in the screening procedure, and were further characterised.

20 **EXAMPLE 5 -Restriction Fragment Length Polymorphism (RFLP) analysis**

There are two genetic loci in *P. hybrida*, Ht1 and Ht2, that control flavonoid 3'-hydroxylase activity (Tabak *et al.*, 1978; Wiering and de Vlaming, 1984). Ht1 is expressed in both the limb and the tube of *P. hybrida* flowers and gives rise to higher levels of F3'H activity than does Ht2 which is only expressed in the tube. The F3'H is able to convert dihydrokaempferol and naringenin to dihydroquercetin and eriodictyol, respectively. In a flower producing delphinidin-based pigments, F3'H activity is masked by the F3'5'H activity. Therefore, the F3'H/F3'5'H assay (Stotz and Forkmann, 1982) is useless in determining the presence or absence of F3'H activity. The enzyme flavonol synthase is able to convert dihydrokaempferol to kaempferol and dihydroquercetin to quercetin (Figure 1a). Myricetin, the 3', 5' hydroxylated flavonol, is produced at low levels in petunia flowers.

- 39 -

Therefore, analysing the flowers for the 3' hydroxylated flavonol, quercetin, infers the presence of F3'H activity.

Restriction Fragment Length Polymorphism (RFLP) analysis of DNA isolated from 5 individual plants in a VR (Ht1/ht1) x V23 (ht1/ht1) backcross was used to determine which, if any, of the cDNA clones representing P450s were linked to the Ht1 locus. Northern analysis of RNA isolated from these plants was also used to detect the presence or absence of a transcript in these lines.

10 Flowers from a VR (Ht1/ht1) x V23 (ht1/ht1) backcross population were analysed for the presence of the flavonols, kaempferol and quercetin. VR (Ht1/ht1) flowers accumulate quercetin and low levels of kaempferol while V23 (ht1/ht1) flowers accumulate kaempferol but little or no quercetin. Individual plants from the VR (Ht1/ht1) x V23 (ht1/ht1) backcross were designated as VR-like (Ht1/ht1), if a substantial level of quercetin was detected in the 15 flower extracts, and V23-like (ht1/ht1), if little or no quercetin but substantial levels of kaempferol were detected in the flower extracts (see Figure 6).

Isolation of Genomic DNA

DNA was isolated from leaf tissue essentially as described by Dellaporta *et al.*, (1983). The 20 DNA preparations were further purified by CsCl buoyant density centrifugation (Sambrook *et al.*, 1989).

Southern blots

The genomic DNA (10 µg) was digested for 16 hours with 60 units of EcoRI and 25 electrophoresed through a 0.7% (w/v) agarose gel in a running buffer of TAE (40 mM Tris-acetate, 50 mM EDTA). The DNA was then denatured in denaturing solution (1.5 M NaCl/0.5 M NaOH) for 1 to 1.5 hours, neutralized in 0.5 M Tris-HCl (pH 7.5)/ 1.5 M NaCl for 2 to 3 hours and then transferred to a Hybond N (Amersham) filter in 20 x SSC.

30 RNA blots

- 40 -

Total RNA was isolated from the petal tissue of *P. hybrida* cv OGR stage 1 to 3 flowers using the method of Turpen and Griffith (1986).

RNA samples were electrophoresed through 2.2 M formaldehyde/1.2% (w/v) agarose gels using running buffer containing 40 mM morpholinopropanesulphonic acid (pH 7.0), 5 mM sodium acetate, 0.1 mM EDTA (pH 8.0). The RNA was transferred to Hybond-N filters (Amersham) as described by the manufacturer.

Hybridization and washing conditions

Southern and RNA blots were probed with ^{32}P -labelled cDNA fragment (10^8 cpm/ μg , 2×10^6 cpm/mL). Prehybridizations (1 hour at 42°C) and hybridizations (16 hours at 42°C) were carried out in 50% (v/v) formamide, 1 M NaCl, 1% (w/v) SDS, 10% (w/v) dextran sulphate. Filters were washed in $2 \times \text{SSC}$, 1% (w/v) SDS at 65°C for 1 to 2 hours and then $0.2 \times \text{SSC}$, 1% (w/v) SDS at 65°C for 0.5 to 1 hour. Filters were exposed to Kodak XAR film with an intensifying screen at -70°C for 16 hours.

RFLP and Northern analysis of the cytochrome P450 fragments

RFLP analysis was used to investigate linkage of the genes corresponding to the OGR-27, OGR-38 and OGR-39 cDNA clones to the *Ht1* locus.

20

^{32}P -labelled fragments of OGR-27, OGR-38 and OGR-39 cDNA clones were used to probe RNA blots and Southern blots of genomic DNA isolated from individual plants in the VR x V23 backcross population. Analysis of *Eco*RI digested genomic DNA isolated from a VR x V23 backcross population revealed a RFLP for the OGR-38 probe which was linked to *Ht1*. Furthermore, a much reduced level of transcript was detected in the V23-like lines, when compared with the high levels of transcript detected in VR-like lines (Figure 6).

25

The data provided strong evidence that the OGR-38 cDNA clone, contained in plasmid pCGP1805, corresponded to the *Ht1* locus and represented a F3'H.

30

- 41 -

RFLP analysis of a V23 x R51 F₂ backcross

RFLP analysis was used to investigate linkage of the gene corresponding to the OGR-38 cDNA to known genetic loci.

- 5 The RFLP linkage analysis was performed using a Macintosh version 2.0 of the MapMaker mapping program (Du Pont) (Lander *et al.*, 1987). A LOD score of 3.0 was used for the linkage threshold.

Analysis of EcoRI or XbaI digested genomic DNA isolated from a V23 x R51 F₂ population
10 revealed a RFLP for the OGR-38 probe which was linked to PAc4. PAc4, a petunia actin cDNA clone (Baird and Meagher, 1987), is a molecular marker for chromosome III and is linked to the HtI locus (McLean *et al.*, 1990). There was co-segregation of the OGR-38 and PAc4 RFLPs for 36 out of 44 V23 x R51 F₂ plants. This represents a recombination frequency of 8% which is similar to a reported recombination frequency of 16% between the
15 Ht1 locus and PAc4 (Cornu *et al.*, 1990).

Further characterisation of OGR-38

The developmental expression profiles in OGR petals, as well as in other OGR tissues, were determined by using the ³²P-labelled fragments of the OGR-38 cDNA insert as a probe
20 against an RNA blot containing 20µg of total RNA isolated from each of the five petunia OGR petal developmental stages as well as from leaves, sepals, roots, stems, peduncles, ovaries, anthers and styles. The OGR-38 probe hybridized with a 1.8kb transcript that peaked at the younger stages of 1 to 3 of flower development. The OGR-38 hybridizing transcript was most abundant in the petals and ovaries and was also detected in the sepals,
25 peduncles and anthers of the OGR plant. A low level of transcript was also detected in the stems. Under the conditions used, no hybridizing transcript was detected by Northern analysis of total RNA isolated from leaf, style or roots.

- 42 -

EXAMPLE 6- Complete sequence of OGR-38

The complete sequence of the OGR-38 cDNA clone (SEQ ID NO:1) was determined by compilation of sequence from different pUC18 subclones obtained using standard procedures for the generation of randomly-overlapping clones (Sambrook *et al.*, 1989). The sequence
5 contained an open reading frame of 1536 bases which encodes a putative polypeptide of 512 amino acids.

The nucleotide and predicted amino acid sequences of OGR-38 (SEQ ID NO:1 and SEQ ID NO:2) were compared with those of the cytochrome P450 probes used in the screening
10 process and with other petunia cytochrome P450 sequences (US Patent Number 5,349,125) using an lfasta alignment (Pearson and Lipman, 1988). The nucleotide sequence of OGR-38 was most similar to the nucleic acid sequence of the flavonoid 3' 5'-hydroxylases representing Hf1 and Hf2 loci from *P. hybrida* (Holton *et al.*, 1993). The Hf1 clone showed 59.6% similarity to the OGR-38 cDNA clone, over 1471 nucleotides, and 49.9%
15 similarity, over 513 amino acids, while the Hf2 clone showed 59.1% similarity to the OGR-38 cDNA clone, over 1481 nucleotides, and 49.0% similarity, over 511 amino acids.

EXAMPLE 7- The F3'H assay of the Hf1 cDNA clone (OGR-38) expressed in yeast**20 Construction of pCGP1646**

The plasmid pCGP1646 (Figure 7) was constructed by cloning the OGR-38 cDNA insert from pCGP1805 in a "sense" orientation behind the yeast glyceraldehyde-3-phosphate dehydrogenase promoter of pYE22m (Tanaka *et al.*, 1988).

25 The plasmid pCGP1805 was linearised by digestion with Asp718. The overhanging 5' ends were "filled in" using DNA polymerase (Klenow fragment) according to standard protocols (Sambrook *et al.*, 1989). The 1.8 kb OGR-38 cDNA fragment was released upon digestion with SmaI. The cDNA fragment was isolated and purified using the Bresaclean kit (Bresatec) and ligated with blunted EcoRI ends of pYE22m. The plasmid pYE22m had been
30 digested with EcoRI and the overhanging 5' ends were removed using DNA polymerase

- 43 -

(Klenow fragment) according to standard protocols (Sambrook *et al.*, 1989). The ligation was carried with the Amersham Ligation kit using 100ng of the 1.8 kb OGR-38 fragment and 150ng of the prepared yeast vector, pYE22m. Correct insertion of the insert in pYE22m was established by XhoI/SaII restriction enzyme analysis of the plasmid DNA isolated from 5 ampicillin-resistant transformants.

Yeast transformation

The yeast strain G-1315 (Mat_α, trp1) (Ashikari *et al.*, 1989) was transformed with pCGP1646 according to Ito *et al.* (1983). The transformants were selected by their ability 10 to restore G-1315 to tryptophan prototrophy.

Preparation of yeast extracts for assay of F3'H activity

A single isolate of G-1315/pCGP1646 was used to inoculate 50 mL of Modified Burkholder's medium (20.0g/L dextrose, 2.0g/L L-asparagine, 1.5g/L KH₂PO₄, 0.5g/L 15 MgSO₄·7H₂O, 0.33g/L CaCl₂, 2g/L (NH₄)₂SO₄, 0.1 mg/L KI, 0.92g/L (NH₄)₆Mo₇O₂₄·4H₂O, 0.1g/L nitrilotriacetic acid, 0.99 mg/L FeSO₄·7H₂O, 1.25 mg/L EDTA, 5.47 mg/L ZnSO₄·7H₂O, 2.5 mg/L FeSO₄·7H₂O, 0.77 mg/L MnSO₄·7H₂O, 0.196 mg/L CuSO₄·5H₂O, 0.124 mg/L Co(NH₄)₂(SO₄)₂·6H₂O, 0.088 mg/L Na₂B₄O₇·10H₂O, 0.2 mg/L thiamine, 0.2 mg/L pyridoxine, 0.2 mg/L nicotinic acid, 0.2 20 mg/L pantothenate, 0.002 mg/L biotin, 10 mg/L inositol) which was subsequently incubated until the value at OD₆₀₀ was 1.8 at 30°C. Cells were collected by centrifugation and resuspended in Buffer 1 [10 mM Tris-HCl buffer (pH 7.5) containing 2 M sorbitol, 0.1 mM DTT, 0.1 mM EDTA, 0.4 mM phenylmethylsulfonyl fluoride (PMSF) and 5 mg yeast lytic enzyme/mL]. Following incubation for 1 hour at 30°C with gentle shaking, the cells were 25 pelleted by centrifugation and washed in ice cold Buffer 2 [10 mM Tris-HCl (pH7.5) containing 0.65 M sorbitol, 0.1 mM DTT, 0.1 mM EDTA, 0.4 mM PMSF]. The cells were then resuspended in Buffer 2 and sonicated using six 15-second bursts with a Branson Sonifier 250 at duty cycle 30% and output control 10%. The sonicated suspension was centrifuged at 10,000 rpm for 30 minutes and the supernatant was centrifuged at 13,000 rpm 30 for 90 minutes. The microsomal pellet was resuspended in assay buffer (100 mM potassium

- 44 -

phosphate (pH 8), 1 mM EDTA, 20 mM 2-mercaptoethanol) and 100 μ L was assayed for activity.

F3'H Assay

5 F3'H enzyme activity was measured using a modified version of the method described by Stotz and Forkmann (1982). The assay reaction mixture typically contained 100 μ L of yeast extract, 5 μ L of 50 mM NADPH in assay buffer (100 mM potassium phosphate (pH8.0), 1 mM EDTA and 20 mM 2-mercaptoethanol) and 10 μ Ci of [3 H]- naringenin and was made up to a final volume of 210 μ L with the assay buffer. Following incubation at 23°C for 2-16
10 hours, the reaction mixture was extracted with 0.5 mL of ethylacetate. The ethylacetate phase was dried under vacuum and then resuspended in 10 μ L of ethylacetate. The tritiated flavonoid molecules were separated on cellulose thin layer plates (Merck Art 5577, Germany) using a chloroform: acetic acid: water (10:9:1 v/v) solvent system. The reaction products were localised by autoradiography and identified by comparison to non-radioactive
15 naringenin and eriodictyol standards which were run alongside the reaction products and visualised under UV light.

F3'H activity was detected in extracts of G1315/pCGP1646, but not in extracts of non-transgenic yeast. From this it was concluded that the cDNA insert from pCGP1805 (OGR-
20 38), which was linked to the *Ht1* locus, encoded a F3'H.

EXAMPLE 8- Transient expression of the *Ht1* cDNA clone (OGR-38) in plants

Construction of pCGP1867

25 Plasmid pCGP1867 (Figure 8) was constructed by cloning the cDNA insert from pCGP1805 in a "sense" orientation behind the Mac promoter (Comai *et al.*, 1990) of pCGP293 (Brugliera *et al.*, 1994). The plasmid pCGP1805 was digested with *Xba*I and *Kpn*I to release the cDNA insert. The cDNA fragment was isolated and purified using the Bresaclean kit (Bresatec) and ligated with *Xba*I/*Kpn*I ends of the pCGP293 binary vector. The ligation
30 was carried out using the Amersham ligation kit. Correct insertion of the fragment in

- 45 -

pCGP1867 was established by ~~XbaI/KpnI~~ restriction enzyme analysis of DNA isolated from gentamycin-resistant transformants.

Transient expression of the **Ht1** cDNA clone (OGR-38) in petunia petals

- 5 In order to rapidly determine whether the OGR-38 cDNA fragment in pCGP1867 represented a functional F3'H in plants, a transient expression study was established. Petals of the mutant *P. hybrida* line Skr4 x SW63 were bombarded with gold particles (1 μ m diameter) coated with pCGP1867 DNA.
- 10 Gold microcarriers were prewashed 3 times in 100% ethanol and resuspended in sterile water. For each shot, 1 μ g of pCGP1867 DNA, 0.5 mg of gold microcarriers, 10 μ L of 2.5 M CaCl₂ and 2 μ L of 100 mM spermidine (free base) were mixed by vortexing for 2 minutes. The DNA coated gold particles were pelleted by centrifugation, washed twice with 100% ethanol and finally resuspended in 10 μ L of 100% ethanol. The suspension was placed
- 15 directly on the centre of the macrocarrier and allowed to dry.

Stages 1 and 2 of Skr4 x SW63 flowers were cut vertically into halves and partially embedded in MS solid media (3% (w/v) sucrose, 100 mg/L myo-inositol, 1xMS salts, 0.5 mg/L pyridoxine-HCl, 0.1 mg/L thiamine-HCl, 0.5 mg/L nicotinic acid and 2 mg/L

20 glycine). The petals were placed so that the inside of the flower buds were facing upwards. A Biolistic PDS-1000/He System (Bio-Rad), using a Helium gas pressure of 900psi and a chamber vacuum of 28 inches of mercury, was used to project the gold microcarriers into the petal tissue. After 6-12 hours under lights in a controlled plant growth room at 22°C, red anthocyanin spots were observed on the upper epidermal layer of the petal tissue

25 bombarded with pCGP1867-coated particles. No coloured spots were observed in control petal bombarded with gold particles alone. These results indicated that the OGR-38 cDNA clone under the control of the Mac promoter was functional, at least transiently, in petal tissue.

**EXAMPLE 9- Stable expression of the *Ht1* cDNA clone (OGR-38) in petunia petals-
Complementation of a *ht1/ht1* petunia cultivar**

5 *A. tumefaciens* transformations

The plasmid pCGP1867 (Figure 8) was introduced into the *Agrobacterium tumefaciens* strain AGL0 by adding 5 µg of plasmid DNA to 100 µL of competent AGL0 cells prepared by inoculating a 50 mL MG/L (Garfinkel and Nester, 1980) culture and growing for 16 hours with shaking at 28°C. The cells were then pelleted and resuspended in 0.5 mL of 85 % (v/v) 10 100 mM CaCl₂/15 % (v/v) glycerol. The DNA-*Agrobacterium* mixture was frozen by incubation in liquid N₂ for 2 minutes and then allowed to thaw by incubation at 37°C for 5 minutes. The DNA/bacterial mix was then placed on ice for a further 10 minutes. The cells were then mixed with 1 mL of LB (Sambrook *et al.*, 1989) media and incubated with shaking for 16 hours at 28°C. Cells of *A. tumefaciens* carrying pCGP1867 were selected 15 on LB agar plates containing 10 µg/mL gentamycin. The presence of pCGP1867 was confirmed by Southern analysis of DNA isolated from the gentamycin-resistant transformants.

Petunia transformations

20 (a) Plant Material

Leaf tissue from mature plants of *P. hybrida* cv Skr4 x SW63 was treated in 1.25 % (w/v) sodium hypochlorite for 2 minutes and then rinsed three times in sterile water. The leaf tissue was then cut into 25 mm² squares and precultured on MS media (Murashige and Skoog, 1962) supplemented with 0.05 mg/L kinetin and 1.0 mg/L 2,4-dichlorophenoxyacetic 25 acid (2,4-D) for 24 hours.

(b) Co-cultivation of *Agrobacterium* and Petunia Tissue

A. tumefaciens strain AGL0 containing the binary vector pCGP1867 (Figure 11) was maintained at 4°C on MG/L agar plates with 100 mg/L gentamycin. A single colony was 30 grown overnight in liquid medium containing 1 % (w/v) Bacto-peptone, 0.5 % (w/v) Bacto-

- 47 -

yeast extract and 1% (w/v) NaCl. A final concentration of 5×10^8 cells/mL was prepared the next day by dilution in liquid MS medium containing B5 vitamins (Gamborg *et al.*, 1968) and 3% (w/v) sucrose (BPM). The leaf discs were dipped for 2 minutes into BPM containing AGL0/pCGP1867. The leaf discs were then blotted dry and placed on co-cultivation media for 4 days. The co-cultivation medium consisted of SH medium (Schenk and Hildebrandt, 1972) supplemented with 0.05 mg/L kinetin and 1.0 mg/L 2,4-D and included a feeder layer of tobacco cell suspension spread over the co-cultivation medium with a filter paper placed on top of the tobacco cell suspension.

10 (c) Recovery of transgenic petunia plants

After co-cultivation, the leaf discs were transferred to selection medium (MS medium supplemented with 3% (w/v) sucrose, α -benzylaminopurine (BAP) 2 mg/L, 0.5 mg/L α -naphthalene acetic acid (NAA), kanamycin 300 mg/L, 350 mg/L cefotaxime and 0.3% (w/v) Gelrite Gellan Gum (Schweizerhall)). Regenerating explants were transferred to fresh selection medium after 4 weeks. Adventitious shoots which survived the kanamycin selection were isolated and transferred to BPM containing 100 mg/L kanamycin and 200 mg/L cefotaxime for root induction. All cultures were maintained under a 16 hour photoperiod (60 μ mol. m⁻², s⁻¹ cool white fluorescent light) at $23 \pm 2^\circ\text{C}$. When roots reached 2-3 cm in length the transgenic petunia plantlets were transferred to autoclaved Debco 51410/2 potting mix in 8 cm tubes. After 4 weeks, plants were replanted into 15 cm pots, using the same potting mix, and maintained at 23°C under a 14 hour photoperiod (300 μ mol. m⁻², s⁻¹ mercury halide light).

25 EXAMPLE 10 -Transgenic plant phenotype analysis pCGP1867 in Skr4 x SW63

Table 5 shows the various petal and pollen colour phenotypes obtained with Skr4 x SW63 plants transformed with the pCGP1867 plasmid. The transgenic plants #593A, 590A, 571A, 589A, 592A and 591A produced flowers with altered petal colour. Moreover, the anthers and pollen of the flowers from plants #593A, 590A, 589A, 592A and 591A were pink,

- 48 -

compared with those of the control Skr4 x SW63 plant, which were white. The change in anther and pollen colour, observed on introduction of plasmid pCGP1867 into Skr4 x SW63 petunia plants, was an unanticipated outcome. The colour codes are taken from the Royal Horticultural Society's Colour Chart (RHSCC). They provide an alternative means by which to describe the colour phenotypes observed. The designated numbers, however, should be taken only as a guide to the perceived colours and should not be regarded as limiting the possible colours which may be obtained.

TABLE 5

Summary of petal, anther and pollen colours obtained in Skr4 xSW63 plants transformed with pCGP1867

Accession Number	Petal Limb Colour	RHSCC Code (petal limb)	Anther & Pollen Colour
Skr4 x SW63 control (594A)	very pale lilac	69B/73D	white
593A	dark pink	67B	pink
590A	dark pink and pink sectors	sectored 67B and 73A	pink
571A	pink	68A and B	pink
589A	dark pink	68A	pink
592A	pink and light pink sectors	68A and 68B	light pink
591A	dark pink	68A	pink
570A	very pale lilac	69B/73D	white

The expression of the introduced Ht1 cDNA in the Skr4 x SW63 hybrid had a marked effect on flower colour. The stamen tissue of the non-transgenic control is white, whereas the same tissue in most of the transgenic plants was pink. In addition, expression of the Ht1 cDNA in the Skr4 x SW63 hybrid conferred a dark pink hue to the corolla, which is normally very pale lilac.

EXAMPLE 11- Analysis of products

The anthocyanidins and flavonols produced in the petals and stamens (included the pollen, anthers and filaments) of the Skr4 x SW63 plants transformed with pCGP1867 were analysed by TLC.

5

Extraction of anthocyanins and flavonols

Prior to TLC analysis, the anthocyanin and flavonol molecules present in petal and stamen extracts were acid hydrolysed to remove glycosyl moieties from the anthocyanidin or flavonol core. Anthocyanidin and flavonol standards were used to help identify the
10 compounds present in the floral extracts.

Anthocyanins and flavonols were extracted and hydrolysed by boiling between 100 to 200 mg of petal limbs, or five stamens, in 1 mL of 2 M hydrochloric acid for 30 minutes. The hydrolysed anthocyanins and flavonols were extracted with 200 μ L of iso-amylalcohol. This
15 mixture was then dried down under vacuum and resuspended in a smaller volume of methanol/1 % (v/v) HCl. The volume of methanol/1 % (v/v) HCl used was based on the initial fresh weight of the petal so that the relative levels of flavonoids in the petals could be estimated. Extracts from the stamens were resuspended in 1 μ L of methanol/1 % (v/v) HCl. A 1 μ L aliquot of the extracts from the pCGP1867 in Skr4 x SW63 petals and stamens was
20 spotted onto a TLC plate.

TLC analysis of floral extracts

Acid-hydrolysed floral extracts were run in a Forestal solvent system (HOAc:water:HCl; 30: 10: 3) (Markham, 1982). Table 6 shows the results of the TLC analysis of the
25 anthocyanidins and flavonols present in some of the flowers and stamens of the transgenic Skr4 x SW63 petunia plants transformed with pCGP1867. Indicative relative amounts of the flavonols and anthocyanidins (designated with a "+" to "+++") were estimated by comparing the intensities of the spots observed on the TLC plate.

30

TABLE 6

Relative levels of anthocyanidins and flavonols detected in the petal limbs and stamens of
Skr4 x SW63 plants transformed with pCGP1867.

5	Acc#	Petal Colour	Anthocyanidins			Flavonols	
			Malvidin	Cyanidin	Peonidin	Kaempferol	Quercetin
	Skr4 x SW63 control petal limb	pale lilac	+/-	-	-	+	-
10	593A petal limb	dark pink	-	+	+++	-	++
	571A petal limb	pink	-	+	+	-	+
	589A petal limb	dark pink	-	+	++	-	++
	570A petal limb	pale lilac	+/-	-	-	+	-
15	Skr4 x SW63 control stamens	white	-	-	-	+++	+
	593A stamens	pink	-	-	++	-	+++

Introduction of the Ht1 cDNA clone into Skr4 x SW63 led to production of the 3'-
20 hydroxylated flavonoids, quercetin, peonidin and some cyanidin in the petals. Peonidin is
the methylated derivative of cyanidin (Figures 1a and 1b). Only kaempferol and a small
amount of malvidin were detected in the non-transgenic Skr4 x SW63 control (Table 6).
Although Skr4 x SW63 is homozygous recessive for both the Hf1 and Hf2 genes, these
mutations do not completely block production of F3'5'H (see US Patent Number 5,349,125)
25 and low levels of malvidin are produced to give the petal limb a pale lilac colour.

The stamens with the pink pollen and anthers produced by the transgenic plant #593A
contained peonidin and quercetin, while the non-transgenic Skr4 x SW63 control with white
pollen and anthers contained kaempferol and a low level of quercetin (Table 6).

30

The accumulation of the 3'-hydroxylated anthocyanidin, peonidin, in the petals and stamens
of the transgenic Skr4 x SW63/pCGP1867 plants correlated with the pink and dark pink
colours observed in the petals, anthers and pollen of the same plants.

Co-suppression of F3'H activity

The plasmid pCGP1867 was also introduced into *P. hybrida* cv. Old Glory Red (Ht1) in order to reduce the level of F3'H activity.

5 Petunia transformations were carried out as described in Example 9, above.

Two out of 38 transgenic plants produced flowers with an altered phenotype. OGR normally produces deep red flowers (RHSCC#46B). The two transgenic plants with altered floral colour produced flowers with a light pink or light red hue (RHSCC#54B and #53C).

10

Northern analysis on RNA isolated from flowers produced by four transgenic plants (the two transgenics with an altered phenotype and two transgenics with the usual deep red flowers) was performed to examine the level of OGR-38 transcripts. Ten micrograms of total petal RNA was separated on a 1.2% (w/v) agarose/formaldehyde gel (Sambrook *et al.* 1989) and
15 transferred to HybondN nylon membrane (Amersham), as described previously. Petal RNA from a non-transformed OGR flower was also included as a control. ³²P-labelled fragments of the OGR-38 cDNA inserts were used to probe the RNA blot.

The OGR-38 probe detected transcripts of approximately 2.4 kb and 1.8 kb in the flowers
20 of the transgenic plants. However, the level of both transcripts detected in the light pink and light red flowers was considerably lower than that detected in the deep red transgenic flowers. The endogenous 1.8 kb transcript was also detected in RNA from the non-transformed OGR flowers. In order to confirm that the 2.4kb transcript was from the introduced OGR-38 transgene, ³²P-labelled fragments of the *mas* terminator region were
25 used to probe the same RNA blot. The *mas* probe detected the 2.4 kb transcript, suggesting that at least this transcript was derived from the introduced OGR-38 transgene.

Analysis of anthocyanin levels

The levels of anthocyanins in the control flowers and in the light pink transgenic flower were
30 measured by spectrophotometric analysis.

- 52 -

Extraction of anthocyanins and flavonols

Anthocyanins and flavonols were extracted from petal limbs by incubating 200 to 300mg of petal limb in 2mL of methanol/1% (v/v) HCl for 16 hours at 4°C. Fifty μ L of this solution was then added to 950 μ L of methanol/1% (v/v) HCl and the absorbance of the diluted solution at 530nm was determined. The anthocyanin level in nmoles per gram was determined using the formula: $[(\text{Abs (530 nm)}/34,000) \times \text{volume of extraction buffer} \times \text{dilution factor} \times 10^6] / \text{weight in grams}$.

The light pink flower was found to contain approximately 915 nmoles of anthocyanin per gram of petal limb tissue whilst the control flower contained around 4000nmol/gram.

These data suggest that introduction of the petunia F3'H (OGR-38) cDNA clone in a sense orientation into OGR plants leads to "co-suppression" (i.e. reduction) of both the endogenous and the transgenic F3'H transcripts. A correlation was observed between lighter flower colours, reduced anthocyanin production and reduced F3'H transcript level.

EXAMPLE 12- Isolation of a F3'H cDNA clone from *Dianthus caryophyllus*

In order to isolate a *Dianthus caryophyllus* (carnation) F3'H cDNA clone, the petunia Ht1-linked F3'H cDNA clone (OGR-38), contained in pCGP1805 (described above), was used to screen a Carnation cv. Kortina Chanel petal cDNA library, under low stringency conditions.

Construction of Carnation cv. Kortina Chanel cDNA library

Twenty micrograms of total RNA isolated (as described previously) from stages 1, 2 and 3 of Kortina Chanel flowers was reverse transcribed in a 50 μ L volume containing 1 x Superscript™ reaction buffer, 10 mM dithiothreitol (DTT), 500 μ M dATP, 500 μ M dGTP, 500 μ M dTTP, 500 μ M 5-methyl-dCTP, 2.8 μ g Primer-Linker oligo from ZAP-cDNA Gigapack III Gold cloning kit (Stratagene) and 2 μ L Superscript™ reverse transcriptase (BRL). The reaction mix was incubated at 37°C for 60 minutes, then placed on ice. A

ZAP-cDNA Gigapack III Gold Cloning kit (Stratagene) was used to complete the library construction. The total number of recombinants was 2.4×10^6 .

A total of 200,000 pfu of the packaged cDNA was plated at 10,000 pfu per 15 cm diameter plate after transfecting XL1-Blue MRF' cells. The plates were incubated at 37°C for 8 hours, then stored overnight at 4°C. Duplicate lifts were taken onto Colony/Plaque Screen™ filters (DuPont) and treated as recommended by the manufacturer.

Screening of Kortina Chanel petal cDNA library for a F3'H cDNA clone

- 10 Prior to hybridization, the duplicate plaque lifts were treated as described previously. The duplicate lifts from the Kortina Chanel petal cDNA library were screened with ^{32}P -labelled fragments of the 1.8 kb EcoRI/XhoI insert from pCGP1805. Low stringency conditions, as described for the screening of the petunia OGR cDNA library, were used.
- 15 One strongly-hybridizing plaque was picked into PSB and rescreened as detailed above to isolate purified plaques. The plasmid contained in the IZAP bacteriophage vector was rescued and named pCGP1807.

The KC-1 cDNA insert contained in pCGP1807 was released upon digestion with
20 EcoRI/XhoI and is around 2 kb. The complete sequence of the KC-1 cDNA clone was determined by compilation of sequence from subclones of the KC-1 cDNA insert. (Partial sequence covering 458 nucleotides had previously been generated from a 800 bp KpnI fragment covering the 3' region of KC-1 which was subcloned into pBluescript to give pCGP1808.) The complete sequence (SEQ ID NO:3) contained an open reading frame of
25 1508 bases which encodes a putative polypeptide of 500 amino acids (SEQ ID NO:4).

The nucleotide and predicted amino acid sequences of the carnation KC-1 cDNA clone were compared with those of the petunia OGR-38 F3'H cDNA clone (SEQ ID NO:1 and SEQ ID NO:2). The sequences of the carnation KC-1 cDNA clone (SEQ ID NO:3 and 4) showed
30 67.3% similarity, over 1555 nucleotides, and 71.5 % similarity, over 488 amino acids, to

- 54 -

that of the petunia OGR-38-F3'H cDNA clone.

An alignment of the petunia, carnation, snapdragon, arabidopsis, rose, chrysanthemum and torenia sequences, all of which are disclosed in this specification, and various summaries of comparisons of sequence similarities among the nucleotide and corresponding amino acid sequences, can be found in Table 7 and in Tables 8, 9, 10, 11 and 12, respectively. These Tables are in Example 34, at the end of the specification.

EXAMPLE 13- Stable expression of the carnation F3'H cDNA (KC-1) clone in petunia petals- Complementation of a ht1/ht1 petunia cultivar

Preparation of pCGP1810

Plasmid pCGP1810 (Figure 9) was constructed by cloning the cDNA insert from pCGP1807 in a "sense" orientation behind the Mac promoter (Comai *et al.*, 1990) of pCGP90 (US Patent Number 5,349,125), a pCGP293 based construct (Brugliera *et al.*, 1994). The plasmid pCGP1807 was digested with BamHI and ApaI to release the KC-1 cDNA insert. The cDNA fragment was isolated and purified using the Bresaclean kit (Bresatec). The pCGP90 binary vector was digested with BamHI and ApaI to release the linearised vector and the HfI cDNA insert. The linearised vector was isolated and purified using the Bresaclean kit (Bresatec) and ligated with BamHI/ApaI ends of the KC-1 cDNA clone. The ligation was carried out using the Amersham ligation. Correct insertion of the insert in pCGP1810 was established by BamHI/ApaI restriction enzyme analysis of DNA isolated from gentamycin-resistant transformants.

The binary vector pCGP1810 was introduced into *A. tumefaciens* strain AGL0 cells, as described in Example 9. The pCGP1810/AGL0 cells were subsequently used to transform Skr4 x SW63 petunia plants (also described in Example 9), to test for stable expression and activity of the enzyme encoded by the gene corresponding to the KC-1 cDNA clone.

EXAMPLE 14-Transgenic plant phenotype analysis

pCGP1810 in Skr4 x SW63

The expression of the introduced KC-1 cDNA in the Skr4 x SW63 hybrid had a marked effect on flower colour. Ten of the twelve transgenic plants transformed with pCGP1810
5 produced flowers with an altered petal colour (RHSCC# 73A), compared with the Skr4 x Sw63 control (RHSCC# 74C). Moreover the anthers and pollen of the transgenic flowers were pink, compared with those of the control Skr4 x SW63 plant, which were white.. In addition, expression of the KC-1 cDNA in the Skr4 x SW63 hybrid conferred a dark pink hue to the corolla, which is normally pale lilac. The colour codes are taken from the Royal
10 Horticultural Society's Colour Chart (RHSCC). They provide an alternative means by which to describe the colour phenotypes observed. The designated numbers, however, should be taken only as a guide to the perceived colours and should not be regarded as limiting the possible colours which may be obtained.

15 Acid-hydrolysed floral extracts (see Example 11) were run in a Forestal solvent system (HOAc:water:HCl; 30: 10: 3) (Markham, 1982). The 3' hydroxylated flavonoids, peonidin and quercetin, were readily detected in the petal limbs of the transgenic plants. Only kaempferol and a small amount of malvidin were detected in the non-transgenic Skr4 x SW63 control.

20

The accumulation of the 3'-hydroxylated anthocyanidin, peonidin, in the petals of the transgenic Skr4 x SW63/pCGP1810 plants correlated with the dark pink colours observed in the petals of the same plants.

25 Construction of pCGP1813

Plasmid pCGP1811 was constructed by cloning the cDNA insert from pCGP1807 in a "sense" orientation behind the Mac promoter (Comai *et al.*, 1990) of pCGP1958. The plasmid pCGP1958 contains the Mac promoter and mannopine synthase (*mas*)(Comai *et al.*, 1990) terminator in a pUC19 backbone. The plasmid pCGP1807 was digested with PstI and
30 XhoI to release the cDNA insert. The overhanging 5' ends were filled in using DNA

- 56 -

polymerase (Klenow fragment) (Sambrook *et al.*, 1989). The cDNA fragment was isolated and purified using the Bresaclean kit (Bresatec) and ligated with SmaI ends of the pCGP1958 vector to produce pCGP1811.

- 5 The plasmid pCGP1811 was subsequently digested with PstI to release the expression cassette containing the Mac promoter driving the KC-1 cDNA with a *mas* terminator, all of which were contained on a 4kb fragment. The expression cassette was isolated and ligated with PstI ends of the pWTT2132 binary vector (DNA Plant Technology Corporation; Oakland, California) to produce pCGP1813 (Figure 10).

10

**Transformation of *Dianthus caryophyllus* cv. Kortina Chanel
with the Carnation F3'H cDNA clone.**

The binary vector pCGP1813 was introduced into *A. tumefaciens* strain AGL0 cells, as described in Example 9. The pCGP1813/AGL0 cells were used to transform carnation
15 plants, to reduce the amount of 3'-hydroxylated flavonoids.

(a) Plant Material

Dianthus caryophyllus (cv. Kortina Chanel) cuttings were obtained from Van Wyk and Son Flower Supply, Victoria, Australia. The outer leaves were removed and the cuttings were
20 sterilised briefly in 70% v/v ethanol followed by 1.25% w/v sodium hypochlorite (with Tween 20) for 6 min and rinsed three times with sterile water. All the visible leaves and axillary buds were removed under the dissecting microscope before co-cultivation.

(b) Co-cultivation of *Agrobacterium* and *Dianthus* Tissue

25 *Agrobacterium tumefaciens* strain AGL0 (Lazo *et al.*, 1991), containing the binary vector pCGP1813, was maintained at 4°C on LB agar plates with 50 mg/L tetracycline. A single colony was grown overnight in liquid LB broth containing 50 mg/L tetracycline and diluted to 5×10^8 cells/mL next day before inoculation. *Dianthus* stem tissue was co-cultivated with *Agrobacterium* for 5 days on MS medium supplemented with 3% w/v sucrose, 0.5
30 mg/L BAP, 0.5 mg/L 2,4-dichlorophenoxy-acetic acid (2,4-D), 100 mM acetosyringone

- 57 -

and 0.25 % w/v Gelrite (pH 5.7).

(c) Recovery of Transgenic *Dianthus* Plants

For selection of transformed stem tissue, the top 6-8 mm of each co-cultivated stem was cut into 3-4 mm segments, which were then transferred to MS medium (Murashige and Skoog, 1962) supplemented with 0.3 % w/v sucrose, 0.5 mg/L BAP, 0.5 mg/L 2,4-D, 1 µg/L chlorsulfuron, 500 mg/L ticarcillin and 0.25 % w/v Gelrite. After 2 weeks, explants were transferred to fresh MS medium containing 3 % sucrose, 0.16 mg/L thidiazuron (TDZ), 0.5 mg/L indole-3-butyric acid (IBA), 2 µg/L chlorsulfuron, 500 mg/L ticarcillin and 0.25 % w/v Gelrite and care was taken at this stage to remove axillary shoots from stem explants. After 3 weeks, healthy adventitious shoots were transferred to hormone free MS medium containing 3 % w/v sucrose, 5 µg/L chlorsulfuron, 500 mg/L ticarcillin, 0.25 % w/v Gelrite. Shoots which survived 5 µg/L chlorsulfuron were transferred to the same medium for shoot elongation.

15

Elongated shoots were transferred to hormone-free MS medium containing 5 µg/L chlorsulfuron, 500 mg/L ticarcillin and 0.4 % w/v Gelrite, in glass jars, for normalisation and root production. All cultures were maintained under a 16 hour photoperiod (120 mE/m²/s cool white fluorescent light) at 23 ± 2°C. Normalised plantlets, approximately 1.5-2 cm tall, were transferred to soil (75 % perlite/25 % peat) for acclimation at 23°C under a 14 hour photoperiod (200 mE/m²/s mercury halide light) for 3-4 weeks. Plants were fertilised with a carnation mix containing 1g/L CaNO₃ and 0.75 g/L of a mixture of microelements plus N:P:K in the ratio 4.7:3.5: 29.2.

25

EXAMPLE 15 -Isolation of a F3'H cDNA clone from *Antirrhinum majus* (Snapdragon) using a differential display approach

A novel approach was employed to isolate a cDNA sequence encoding F3'H from *Antirrhinum majus* (snapdragon). Modified methods based on the protocols for (i) isolation of plant cytochrome P450 sequences using redundant oligonucleotides (Holton *et al.* 1993)

30

- 58 -

and (ii) differential display of eukaryotic messenger RNA (Liang and Pardee, 1992) were combined, to compare the petal cytochrome P450 transcript populations between wild type (Eos^+) and F3'H mutant (eos^-) snapdragon lines. Direct cloning of differentially expressed cDNA fragments allowed their further characterisation by Northern, RFLP and sequence analysis to identify putative F3'H encoding sequences. A full-length cDNA was obtained using the RACE protocol of Frohman *et al.* (1988) and the clone was shown to encode a functional F3'H following both transient and stable expression in petunia petal cells.

Plant Material

- 10 The *Antirrhinum majus* lines used were derived from the parental lines K16 (eos^-) and N8 (Eos^+). K16 is a homozygous recessive mutant lacking F3'H activity, while N8 is wild type for F3'H activity. These lines are similar, though not isogenic. The seed of capsule E228² from the selfed K16 x N8 F₁ plant (#E228) was germinated and the resultant plants (K16 x N8 F₂ plants) were scored for the presence or absence of cyanidin, a product of F3'H activity (see Figures 1a and 1b). The presence of cyanidin could be scored visually, as the flowers were a crimson colour, unlike the mutant plants which were a pink colour (from pelargonidin-derived pigments). The accuracy of the visual scoring was confirmed by TLC analysis of the petal anthocyanins, carried out as described in Example 11.
- 20 Of 13 plants raised from the E228² seed, 9 (#3, #4, #5, #6, #7, #9, #10, #12, #15) produced flowers with cyanidin (Eos^+/Eos^+ and Eos^+/eos^-) while 4 (#8, #11, #13, #14) produced only pelargonidin-derived pigments (eos^-/eos^-).

Synthesis of cDNA

- 25 Total RNA was isolated from the leaves of plant #13 and petal tissue of plants #3, #5, and #12 of the *A. majus* K16 x N8 F₂ segregating population (E228²) using the method of Turpen and Griffith (1986). Contaminating DNA was removed by treating 50 μ g total RNA with 1 unit RQ1 RNase-free DNase (Promega) in the presence of 40 units RNasin[®] ribonuclease inhibitor (Promega) for 3 hours at 37°C in a buffer recommended by the manufacturers. The RNA was then further purified by extraction with phenol/chloroform/iso-
- 30

- 59 -

amyl alcohol (25:24:1) and subsequent ethanol precipitation.

Anchored poly(T) oligonucleotides, complementary to the upstream region of the polyadenylation sequence, were used to prime cDNA synthesis from *A. majus* petal and leaf RNA. The oligonucleotide sequences synthesized were (5'-3'):

polyT-anchA	TTTTTTTTTTTTTTTTTA	SEQ ID NO:27
polyT-anchC	TTTTTTTTTTTTTTTTTC	SEQ ID NO:28
polyT-anchG	TTTTTTTTTTTTTTTTTG	SEQ ID NO:29

10

Two micrograms of total RNA and 100 pmol of the appropriate priming oligonucleotide were heated to 70°C for 10 minutes, then chilled on ice. The RNA/primer hybrids were then added to a reaction containing 20 units RNasin® (Promega), 25 nM each dNTP, 10 mM DTT and 1x Superscript buffer (BRL). This reaction was heated at 37°C for 2 minutes, then 200 units of Superscript™ reverse transcriptase (BRL) were added and the reaction allowed to proceed for 75 minutes, after which the reverse transcriptase was inactivated by heating the mixture at 95°C for 20 minutes.

Amplification of cytochrome P450 sequences using PCR

Cytochrome P450 sequences were amplified using redundant oligonucleotides (designed to be complementary to conserved regions near the 3' end of plant cytochrome P450 coding sequences) and polyT anchored oligonucleotides. A similar approach was previously used to generate cytochrome P450 sequences from *Petunia hybrida* and is described in US Patent Number 5,349,125.

25

Four oligonucleotides (referred to as upstream primers) were synthesized. These were derived from conserved amino acid regions in plant cytochrome P450 sequences. The oligonucleotides (written 5' to 3') were as follows:

30 WAIGRDP TGG GCI ATI GGI (A/C)GI GA(T/C) CC

- 60 -

 SEQ ID NO:30 SEQ ID NO:31

FRPERF AGG AAT T(T/C)(A/C) GIC CIG A(A/G)(A/C) GIT T

SEQ ID NO:32 SEQ ID NO:33

5

Pet Haem-New CCI TT(T/C) GGI GCI GGI (A/C)GI (A/C)GI ATI TG(T/G)
(C/G)CI GG

SEQ ID NO:34

10 EFXPERF GAI TT(T/C) III CCI GAI (A/C)GI TT

SEQ ID NO:35 SEQ ID NO:36

The upstream primers were used with each of the polyT anchored oligonucleotides to generate cytochrome P450 sequences in polymerase chain reactions using cDNA as a

15 template. Fifty pmol of each oligonucleotide was combined with 2 μ M of each dNTP, 1.5 mM MgCl₂, 1x PCR buffer (Perkin Elmer), 5 μ Ci α -[³²P] dATP (Bresatec, 1500 Ci/mmol), 2.5 units AmpliTaq® DNA polymerase (Perkin Elmer) and 1/10th of the polyT-anchor primed cDNA reaction (from above). Reaction mixes (50 μ L) were cycled 40 times between 94°C for 15 seconds, 42°C for 15 seconds, and 70°C for 45 seconds, following an initial

20 2 minute denaturation step at 94°C. Cycling reactions were performed using a Perkin Elmer 9600 Gene Amp Thermal Cycler.

DNA sequences were amplified using each upstream primer/anchored primer combination and the appropriately-primed cDNA template. Each primer combination was used with the

25 cDNA from the petals of the E228² plants #3 and #5 (cyanidin-producing flowers) and #12 (non-cyanidin producing flowers). Reactions incorporating leaf cDNA from plant #13 (cyanidin-producing flowers) were also included, as negative controls, because F3'H activity is not present at a significant level in healthy, unstressed leaf tissues.

30 **Differential display of cytochrome P450 sequences**

~~³²P-labelled PCR fragments were visualised following separation on a 5% (w/v)~~
polyacrylamide/urea denaturing gel (Sambrook *et al.* 1989). A ³²P-labelled M13mp18
sequencing ladder was included on the gel to serve as a size marker. The sequencing gel was
dried onto Whatman 3MM paper and exposed to Kodak XAR film at room temperature.

5

Comparison of bands between cyanidin-producing petal samples and the non-cyanidin petal
sample revealed 11 bands which represent mRNAs exclusively present in the cyanidin-
producing petals. Of these 11 bands, only two were also present (at a reduced intensity) in
the leaf sample.

10

Isolation and cloning of PCR fragments from sequencing gel

PCR products were purified from the dried sequencing gel and reamplified by the method
described by Liang *et al.* (1993). Amplified cDNAs were purified, following electrophoretic
separation on a 1.2% (w/v) agarose/TAE gel, using a Bresaclean kit (Bresatec). The purified
15 fragments were then directly ligated into either commercially-prepared pCR-Script™ vector
(Stratagene) or *EcoRV*-linearised pBluescript® (Stratagene) which had been T-tailed using
the protocol of Marchuk *et al.* (1990).

Sequence of F3'H PCR products

20 Each of the eleven cloned differential display PCR products (with inserts not exceeding 500
bp) was sequenced on both strands and compared to other known cytochrome P450
sequences involved in anthocyanin biosynthesis, using the FASTA algorithm of Pearson and
Lipman (1988).

25 Of the eleven cDNAs cloned, two (Am1Gb and Am3Ga) displayed strong homology with
the petunia OGR-38 F3'H sequence (Examples 4 to 11) and the F3'5'H sequences (Holton
et al., 1993). Conserved sequences between clones Am1Gb and Am3Ga suggested that they
represented overlapping fragments of the same mRNA. Clone Am3Ga extends from the
sequence encoding the haem-binding region of the molecule (as recognised by the "Pet
30 Haem-New" oligonucleotide; SEQ ID NO:34) to the polyadenylation sequence. Clone

- 62 -

Am1Gb extends from the cytochrome P450 sequence encoding the conserved "WAIGRDP" amino acid motif (complementary to primer 1; SEQ ID NO:30 and SEQ ID NO:31) to an area in the 3' untranslated region which was spuriously recognised by the primer 1 ("WAIGRDP") oligonucleotide.

5

EXAMPLE 16- RFLP analysis of cytochrome P450 cDNAs

Restriction fragment length polymorphism (RFLP) analysis was again used to investigate linkage of the gene corresponding to cDNA clone Am3Ga to the presence, or absence, of cyanidin-producing activity in petals. A ^{32}P -labelled insert of Am3Ga was used to probe Southern blots of genomic DNA isolated from K16 x N8 F₂ segregating plants as well as the parental K16 and N8 lines. Analysis of *EcoRV*-digested genomic DNA from 13 plants of the K16 x N8 F₂ segregating population revealed hybridization only to the sequences of N8 and the K16 x N8 F₂ segregating lines which displayed floral cyanidin production (Figure 11). The K16 x N8 F₂ plants which produced only pelargonidin-derived pigments in their petals (including parental line, K16) showed no specific hybridization (Figure 11, lanes 2, 8, 11, 13, 14). These data indicate a possible deletion of the genomic sequences corresponding to Am3Ga in the mutant K16 plant and, therefore, at least a partial deletion of the F3'H gene in this line.

20

EXAMPLE 17- Northern analysis of cytochrome P450 cDNAs

Northern analysis was used to confirm the expression profiles of the putative cytochrome P450 fragments as shown by differential display. Ten micrograms of total petal RNA from eight of the K16 x N8 F₂ segregating population was separated on a 1.2% (w/v) agarose/formaldehyde gel (Sambrook *et al.* 1989) and transferred to HybondN nylon membrane (Amersham). Leaf RNA from the cyanidin-producing plant #13 was also included as a negative control in the Northern analysis. ^{32}P -labelled fragments of the cDNA insert from clone Am3Ga was used to probe the RNA blot.

30

- 63 -

The Am3Ga probe recognised an approximately 1.8 kb transcript which was only detectable in the petals of cyanidin-producing plants (plants #1, #3, #4, #5, #8). No transcript was detectable in the pelargonidin-producing petals (plants #6, #11, #12) or in the leaf sample from plant #13 (Figure 12).

5

These data, taken with those of the RFLP analysis, provide strong evidence that Am3Ga clone represents a cytochrome P450 gene which is responsible for F3'H activity in snapdragon. The total lack of a detectable transcript in the petals of non-cyanidin-producing lines supports the findings of the RFLP analysis, that the loss of cyanidin-producing activity in the K16 line (and the homozygous recessive plants of the K16 x N8 F₂ segregating population) is the result of a deletion in the F3'H structural gene.

EXAMPLE 18- Isolation of a full-length snapdragon F3'H cDNA

15 The Rapid Amplification of cDNA Ends (RACE) protocol of Frohman *et al.* (1988) was employed to isolate a full-length F3'H cDNA clone using sequence knowledge of the partial Am3Ga clone. A gene-specific primer ("SnapredRace A" -complementary to Am3Ga sequences 361 to 334) was synthesized to allow reverse transcription from petal RNA. A 3' amplification primer ("SnapredRace C" -complementary to Am3Ga (3'UTR) sequences 283 to 259) was also synthesized to bind just upstream of "SnapredRace A". A "poly(C)" primer was used to amplify sequences from the 5' end of the cDNA molecule.

The sequences of the oligonucleotides used were (written 5'-3'):

25 Snapred Race A CCA CAC GAG TAG TTT TGG CAT TTG ACC C
SEQ ID NO:37

Snapred Race C GTC TTG GAC ATC ACA CTT CAA TCT G
SEQ ID NO:38

30

- 64 -

PolyC race CCG AAT TCC CCC CCC CC

SEQ ID NO:39

"Snapred Race A-primed" petal cDNA was poly(G)-tailed and a 5' cDNA fragment
 5 amplified with primers "Snapred Race C" and "PolyC race" using the method of Frohman
et al. (1988). Pfu DNA polymerase (0.15 unit) (Stratagene) was combined with 2.5 units
 AmpliTaq® DNA polymerase (Perkin Elmer) to increase the fidelity of the PCR reaction.

The resultant 1.71 kb DNA fragment (sdF3'H) was cloned directly into EcoRV-linearised
 10 pBluescript® (Stratagene) vector which had been T-tailed using the protocol of Marchuk *et al.* (1990). This plasmid was named pCGP246.

EXAMPLE 19- Complete sequence of snapdragon F3'H

15 Convenient restriction sites within the sdF3'H cDNA sequence of pCGP246 were exploited
 to generate a series of short overlapping subclones in the plasmid vector pUC19. The
 sequence of each of these subclones was compiled to provide the sequence of the entire
 sdF3'H RACE cDNA. The sdF3'H cDNA sequence was coupled with that from clone
 Am3Ga to provide the entire sequence of a snapdragon F3'H cDNA (SEQ ID NO:5). It
 20 contains an open reading frame of 1711 bases which encodes a putative polypeptide of 512
 amino acids (SEQ ID NO:6).

The nucleotide and predicted amino acid sequences of the snapdragon sdF3'H clone were
 compared with: those of the petunia OGR-38 cDNA clone (SEQ ID NO:1 and SEQ ID
 25 NO:2); the petunia F3'5'H cDNA clones Hf1 and Hf2; and other petunia cytochrome P450
 sequences isolated previously (US Patent Number 5,349,125). The sequence of sdF3'H was
 most similar to that of the petunia F3'H cDNA clone (OGR-38) representing the Ht1 locus
 from *P. hybrida*, having 69% similarity at the nucleic acid level, over 1573 nucleotides, and
 72.2% similarity at the amino acid level, over 507 amino acids.

- 65 -

The Hf1 clone showed 57.3% similarity, over 1563 nucleotides and 49.3% similarity, over 491 amino acids, to the snapdragon sdF3'H clone, while the Hf2 clone showed 57.7% similarity, over 1488 nucleotides, and 50.8% similarity, over 508 amino acids, to the snapdragon sdF3'H clone.

5

The snapdragon sdF3'H sequence contains two "in frame" ATG codons which could act to initiate translation. Initiation from the first of these codons (position 91 of SEQ ID NO:5) gives a protein with an additional 10 N-terminal amino acids and would be favoured according to the scanning model for translation (Kozak, 1989).

10

An alignment of the petunia, carnation, snapdragon, arabidopsis, rose, chrysanthemum and torenia sequences, all of which are disclosed in this specification, and various summaries of comparisons of sequence similarities among the nucleotide and corresponding amino acid sequences, can be found in Table 7 and in Tables 8, 9, 10, 11 and 12, respectively. These

15 Tables are in Example 34, at the end of the specification.

EXAMPLE 20- Transient expression of sdF3'H in plants

Construction of pCGP250

Plasmid pCGP250 (Figure 13) was created by cloning the entire sdF3'H RACE cDNA insert
20 (from position 1 to 1711 (SEQ ID NO:5)) from pCGP246 in the "sense" orientation behind the Mac promoter (Comai *et al.*, 1990) of pCGP293 (Brugliera *et al.*, 1994). The plasmid pCGP246 was digested with EcoRI to release the cDNA insert. The cDNA fragment was blunt-ended by repairing the overhangs with the Klenow fragment of DNA polymerase I (Sambrook *et al.*, 1989) and purified, following agarose gel electrophoresis, using a
25 Bresaclean kit (Bresatec). The blunt cDNA fragment was then ligated into the binary vector pCGP293, which had been linearised with XbaI and blunt-ended using the Klenow fragment of DNA polymerase I. The ligation was carried out using the Amersham ligation kit. Correct insertion of the insert in pCGP250 was established by BamHI and PstI restriction enzyme analysis of DNA isolated from gentamycin-resistant transformants.

30

Construction of pCGP231

Plasmid pCGP231 (Figure 14) was created by cloning the RACE cDNA insert from pCGP246, downstream of the first "in-frame" ATG codon (from position 104 to 1711 (SEQ ID NO:5), in the "sense" orientation behind the Mac promoter (Comai *et al.*, 1990) of pCGP293 (Brugliera *et al.*, 1994). The plasmid pCGP246 was digested with SspI (which recognises a site between the candidate ATG codons) and SmaI (with a site in the vector polylinker sequence) to release a blunt-ended cDNA fragment which includes the entire coding region downstream from the second putative initiation codon. The cDNA fragment was then ligated into the binary vector pCGP293, which had been linearised with XbaI and blunt-ended using the Klenow fragment of DNA polymerase I. The ligation was carried out using the Amersham ligation kit. Correct insertion of the insert in pCGP231 was established by BamHI and PstI restriction enzyme analysis of DNA isolated from gentamycin-resistant transformants.

15 Transient Expression Studies

To determine rapidly whether the pCGP246 sequences in pCGP231 and pCGP250 encoded active flavonoid 3'-hydroxylases in plants, a transient expression study was undertaken. Petals of the mutant *P. hybrida* line Skr4 X SW63 were bombarded with gold particles (1µm diameter) coated with either pCGP231 or pCGP250 plasmid DNA, using the method described in Example 8.

After 6-12 hours under lights in a controlled plant growth room at 22°C, red anthocyanin spots were observed on the surface of the petal tissue bombarded with pCGP231 coated particles. No coloured spots were observed in petals bombarded with pCGP250 or control petals bombarded with gold particles alone. These results indicated that the pCGP246 coding region (starting at the second ATG, position 121 of SEQ ID NO:5), under the control of the Mac promoter, was functional in petal tissue.

EXAMPLE 21- Stable expression of the snapdragon F3'H cDNA clone in petunia petals-
30 **Complementation of a ht1/ht1 petunia cultivar**

- 67 -

The binary vectors pCGP250 and pCGP231 were introduced into *A. tumefaciens* strain AGL0 cells, as described in Example 9. The pCGP250/AGL0 and pCGP231/AGL0 cells were used to transform Skr4 x SW63 petunia plants (also described in Example 9), to test for stable expression and activity of the enzyme encoded by the gene corresponding to the snapdragon sdF3'H cDNA clone.

Three of the nine transgenic plants transformed with pCGP250 produced flowers with a slightly-altered petal colour (RHSCC# 73A), compared with the Skr4 x Sw63 control (RHSCC# 75C). Of the 11 transgenic plants transformed with pCGP231, one plant produced flowers with an altered petal colour (RHSCC# 73B). The anthers and pollen of the transgenic flowers were also white, as in the control. The codes are taken from the Royal Horticultural Society's Colour Chart (RHSCC). They provide an alternative means by which to describe the colour phenotypes observed. The designated numbers, however, should be taken only as a guide to the perceived colours and should not be regarded as limiting the possible colours which may be obtained.

TLC analysis of floral extracts

Acid-hydrolysed floral extracts (see Example 11) were run in a Forestal solvent system (HOAc:water:HCl; 30: 10: 3) (Markham, 1982). Introduction of the sdF3'H cDNA clone into Skr4 x SW63 led to the production of increased levels of the 3'-hydroxylated flavonoid, peonidin, in the petals. Peonidin is the methylated derivative of cyanidin (Figures 1a and 1b).

25 **EXAMPLE 22- Isolation of a F3'H cDNA clone from *Arabidopsis thaliana* using a PCR approach**

In order to isolate a cDNA clone representing flavonoid 3'-hydroxylase from *Arabidopsis thaliana*, PCR fragments were generated using primers from the conserved regions of cytochrome P450s. One PCR product (p58092.13) was found to have high sequence similarity with the petunia OGR-38 and snapdragon F3'H cDNA clones. The PCR fragment

- 68 -

was then used, together with the Ht1 cDNA insert (OGR-38) from pCGP1805, to screen an *A. thaliana* cDNA library.

Design of oligonucleotides

- 5 Degenerate oligonucleotides for PCR DNA amplification were designed from the consensus amino acid sequence of *Petunia hybrida* cytochrome P450 partial sequences situated near the haem-binding domain. Primer degeneracy was established by the inclusion of deoxyinosine (designated as I below) in the third base of each codon (deoxyinosine base pairs with similar efficiency to A, T, G, and C), and the inclusion of alternate bases where the
- 10 consensus sequences were non-specific. Thus, the amino-terminal directional primer "Pet Haem" (*Petunia* haem-binding domain), containing the cysteine residue codon crucial for haem binding, and the upstream primer "WAIGRDP" (See also Example 15) were designed.

WAIGRDP TGG GCI ATI GGI (A/C)GI GA(T/C) CC
15 SEQ ID NO:30 SEQ ID NO:31

Pet Haem CCI GG(A/G) CAI ATI C(G/T)(C/T) (C/T)TI CCI GCI CC(A/G) AAI GG
SEQ ID NO:40

20 Generation of cytochrome P450 sequences using PCR

Genomic DNA was isolated from *A. thaliana* ecotype Columbia, using the method described by Dellaporta *et al.* (1987). Polymerase chain reactions for amplification of cytochrome P450 homologues typically contained 100-200 ng of Columbia genomic DNA, 10 mM Tris-HCl (pH8.3), 50 mM KCl, 1.5 mM MgCl₂, 0.01 % (w/v) gelatin, 0.2 mM each dNTP, 312

25 ng "WAIGRDP" and 484 ng "Pet Haem" and 1.25 units Taq polymerase (Cetus). Reaction mixes (50 µL) were cycled 40 times between 95°C for 50 seconds, 45°C for 50 seconds and 72°C for 45 seconds.

The expected size of specific PCR amplification products, using the "WAIGRDP" and "Pet

30 Haem" primers on a typical P450 gene template, without an intron, is approximately 150

- 69 -

base pairs. PCR fragments of approximately 140 to 155 base pairs were isolated and purified using the Mermaid® kit (BIO 101). The PCR fragments were re-amplified to obtain enough product for cloning and then end-repaired using Pfu DNA polymerase and finally cloned into pCR-Script™Direct SK(+) (Stratagene). The ligated DNA was then used to transform 5 competent DH5α cells (Inoue *et al.*, 1990).

Sequence of PCR products

Plasmid DNA from 15 transformants was prepared (Del Sal *et al.*, 1989). Sequencing data generated from these PCR fragments indicated that 11 out of the 15 represented unique 10 clones. A distinct set of cytochrome P450 consensus amino acids was also found in the translated sequence encoded within the *A. thaliana* PCR inserts. The sequences of the PCR fragments were also compared with those of the petunia OGR-38 F3'H cDNA clone and the snapdragon F3'H cDNA clone. The PCR fragment, p58092.13, was most similar to the F3'H sequences from both petunia and snapdragon.

15

EXAMPLE 23- Screening of *A. thaliana* cDNA library

To isolate a cDNA clone of the p58092.13 PCR product, an *A. thaliana* ecotype Columbia cDNA library (Newman *et al.*, 1994; D' Alessio *et al.*, 1992) was screened with a ³²P- 20 labelled fragment of p58092.13 together with a ³²P-labelled fragment of the petunia Ht1 cDNA insert (OGR-38), contained in pCGP1805.

A total of 600,000 pfu was plated at a density of 50,000 pfus per 15 cm diameter plate, as described by D' Alessio *et al* (1992). After phage growth at 37°C plates were stored at 4°C 25 overnight, duplicate lifts were taken onto Colony/Plaque Screen™ filters (DuPont) and treated as recommended by the manufacturer.

Prior to hybridization, the duplicate plaque lifts were washed in prewashing solution (50 mM Tris-HCl pH7.5, 1 M NaCl, 1 mM EDTA, 0.1% (w/v) sarcosine) at 65°C for 30 minutes; 30 stripped in 0.4 M sodium hydroxide at 65°C for 30 minutes; then washed in a solution of

- 70 -

0.2 M Tris-HCl pH 8.0, 0.1 x SSC, 0.1% (w/v) SDS at 65°C for 30 minutes and finally rinsed in 2 x SSC, 1.0% (w/v) SDS.

Hybridization conditions included a prehybridization step in 50% (v/v) formamide, 1 M NaCl, 10% (w/v) dextran sulphate, 1% (w/v) SDS at 42°C for at least 1 hour. The ³²P-labelled fragment of p58092.13 (2x10⁶cpm/mL) was then added to the hybridization solution and hybridization was continued at 42°C for a further 16 hours. The filters were then washed in 2 x SSC, 1% (w/v) SDS at 42°C for 2 x 1 hour and exposed to Kodak XAR film with an intensifying screen at -70°C for 16 hours.

10

Eleven strongly-hybridizing plaques were picked into PSB and rescreened as detailed above, to isolate purified plaques. These filters were also probed with ³²P-labelled fragment of the petunia *Ht1* cDNA insert (OGR-38), contained in pCGP1805, under low stringency conditions. Low stringency conditions included prehybridization and hybridization at 42°C in 20% (v/v) formamide, 1 M NaCl, 10% (w/v) dextran sulphate, 1% (w/v) SDS and washing in 6xSSC, 1% (w/v) SDS (w/v) at 65°C for 1 hour.

The OGR-38 and p58092.13 probes hybridized with identical plaques. The 11 pure plaques were picked into PSB and the plasmid vectors pZL1 containing the cDNA clones were rescued using the bacterial strain DH10B(Zip). Plasmid DNA was prepared (Del Sal *et al.*, 1989) and the cDNA inserts were released upon digestion with *Bam*HI and *Eco*RI. The 11 plasmids contained cDNA inserts of between 800bp and 1 kb. Sequence data generated from the 5' region of the cDNA inserts suggested that nine of these clones were identical. Sequence data were generated from the 5' ends of all nine cDNA inserts and the 3' end of only one cDNA insert. The sequence data generated from all clones were compiled to produce the nucleotide and translated sequence shown as SEQ ID NO:7 and SEQ ID NO:8.

The *A. thaliana* putative F3'H sequences were compared with the sequences of the petunia OGR-38 F3'H cDNA clone (SEQ ID NO:1 and SEQ ID NO:2) and was 64.7% similar to the petunia F3'H cDNA clone, over 745 nucleotides, and 63.7 % similar, over 248 amino

30

- 71 -

acids.

An alignment of the petunia, carnation, snapdragon, arabidopsis, rose, chrysanthemum and torenia sequences, all of which are disclosed in this specification, and various summaries of comparisons of sequence similarities among the nucleotide and corresponding amino acid sequences, can be found in Table 7 and in Tables 8, 9, 10, 11 and 12, respectively. These Tables are in Example 34, at the end of the specification.

Isolation of a F3'H genomic clone from *Arabidopsis thaliana*

10 To isolate a genomic clone of the *A. thaliana* F3'H gene, a *A. thaliana* ecotype Landsberg erecta genomic DNA library was screened with ³²P-labelled p60606.04 fragments. The library was created by cloning partial MboI-digested genomic DNA between BamHI-digested bacteriophage lambda EMBL4 arms. The primary library, which contained 30,000 clones, was amplified once before screening.

15

The p60606.04 clone, containing a 1 kb fragment of *A. thaliana* F3'H cDNA, was digested with BamHI/EcoRI to excise the insert which was purified using GeneClean (Bio 101). Probe was ³²P-labelled using the nick-translation procedure (Sambrook et al., 1989). Approximately 20,000 plaques were probed at high stringency (50% formamide at 37° C) and filters were washed in: 2x SSPE; 2x SSPE, 0.1% (w/v) SDS; 0.1x SSPE, all at 65°C. Re-screening was carried out under the same conditions.

DNA was purified from three positive plaques (λ TT7-1, λ TT7-5 and λ TT7-6) and mapped by digestion with EcoRI and EcoRI/SalI. All three clones had an EcoRI fragment in common. λ TT7-1 and λ TT7-5 had overlapping but not identical restriction patterns. A Southern blot of these digests was probed as above and, for λ TT7-1 and λ TT7-5, a common 6.5 kb EcoRI/SalI fragment hybridized. A smaller EcoRI/SalI fragment in λ TT7-6 also hybridized and was presumably at the insert boundary.

30 EcoRI/SalI fragments from ITT7-5 were cloned into pBlueScript SK+ and a clone containing

- 72 -

the 6.5 kb fragment, designated E-5, was identified by hybridization (as above) and insert size. A restriction map was compiled for the fragment using EcoRI, SalI, KpnI, HindIII and BglII in various combinations, and by hybridization to Southern blots of these digests with the BamHI/EcoRI insert from the *A. thaliana* F3'H cDNA clone.

5

Complete sequence of Tt7 genomic clone

A 6.4 kb BamHI fragment from pTt7-2, containing most of the Tt7 genomic fragment was purified, self-ligated, sonicated, end-repaired, size-fractionated (450bp to 800bp) and cloned into SmaI-cut pUC19 using standard techniques (Sambrook et al., 1989). Recombinant clones were isolated, and plasmid DNA was purified and sequenced using M13-21 or M13 reverse sequencing primers. The sequence from overlapping clones was combined into one contiguous fragment. The sequence of the ends of the Tt7 genomic fragment were also obtained by sequencing with the -21 and REV primers. All of the sequences were combined together to obtain the complete sequence of the 6.5 kb EcoRI/SalI fragment from E-5 (SEQ ID NO:9).

The sequences over the coding region of the arabidopsis Tt7 genomic clone (SEQ ID NO:10, 11, 12 and 13) were compared with those of the petunia OGR-38 F3'H cDNA clone (SEQ ID NO:1 and 2). The arabidopsis Tt7 coding region showed 65.4% similarity, over 1066 nucleotides, and 67.1% similarity, over 511 amino acids, to that of the petunia OGR-38 F3'H cDNA clone.

Transformation of a tt7 Arabidopsis mutant

Preparation of binary vector

The EcoRI/SalI fragment from E-5 was cloned into EcoRI/SalI-cut pBI101 (Jefferson *et al.*, 1987). Two separate but identical clones were identified: pBI-Tt7-2 (Figure 15) and pBI-Tt7-4. Both clones were used for transformation of *A. tumefaciens*.

Plant Transformation

Plasmids pBI-Tt7-2, pBI-Tt7-4 and pBI101 were transformed into *Agrobacterium* strain

- 73 -

GV3101-pMP90 by electroporation. Transformants were selected on medium containing 50 $\mu\text{g/mL}$ kanamycin (and 50 $\mu\text{g/mL}$ gentamycin to select for the resident pMP90).

Plasmid DNA, from four transformant colonies for each clone, was isolated and digested with EcoRI/SalI, electrophoresed, Southern blotted, and probed with the Tt7 cDNA insert. For pBI-Tt7-2 and pBI-Tt7-4, the expected insert band was identified.

One transformant for each plasmid (i.e.: one control [pBI101 C4], one each of the two Tt7 clones [pBI-Tt7-2-3 and pBI Tt7-4-4]) was used to vacuum infiltrate the *A. thaliana* tt7 mutant line NW88 (4 pots of 10 plants each for each construct), using the a method essentially as described by Bechtold *et al.* (1993).

Seed from each pot was harvested. One hundred mg of seed (approximately 5,000) was plated on nutrient medium (described by Haughn and Somerville, 1986) containing 50 $\mu\text{g/mL}$ kanamycin. Kanamycin-resistant transformants were visible after 7 to 10 days. In the case of pBI-Tt7-2-3 and pBI-Tt7-4-4, a total of 11 transformants were isolated from 5 different seed lots (i.e.: pots) and all kanamycin-resistant transformants were visibly Tt7 in phenotype and exhibited the characteristic red/purple anthocyanin pigments at the margins of the cotyledons and at the hypocotyl. A single kanamycin-resistant transformant was isolated from only one of the four pots of control transformants and it did not exhibit a "wild-type" Tt7 phenotype.

Complementation of tt7 mutant

These transformants were planted out and grown to maturity and individually harvested for seed. In each case, for pBI-Tt7-2-3 and pBI-Tt7-4-4 transformants, the seeds were visibly more brown than the pale brown seed of the tt7 mutant plants. The seed from the control transformant was indistinguishable from the tt7 mutant parent. These seed were plated out on nutrient medium and nutrient medium with kanamycin added, and scored for the Tt7 phenotype (red/purple anthocyanin pigments at the margins of the cotyledons and at the hypocotyl) and kanamycin resistance. The progeny of at least one transformant for each seed

- 74 -

~~lot was examined, since these were clearly independent transformation events.~~

Without exception, kanamycin-resistant seedlings exhibited the *Tt7* phenotype while kanamycin-sensitive individuals were *tt7*. In some cases, kanamycin resistance was weak
5 and variable among a family of seed and it was difficult to unequivocally determine whether individuals were kanamycin resistant or kanamycin sensitive.

EXAMPLE 24- Isolation of a F3'H cDNA clone from *Rosa hybrida*

10 In order to isolate a Rose F3'H cDNA clone, a *Rosa hybrida* cv. Kardinal petal cDNA library was screened with ³²P-labelled fragments of the petunia *Ht1* cDNA clone (OGR-38), contained in pCGP1805, and snapdragon F3'H cDNA clone (sdF3'H), contained in pCGP246.

15 Construction of a petal cDNA library from Rose cv. Kardinal

Total RNA was prepared from the buds of *Rosa hybrida* cv. Kardinal stage 2. At this stage, the tightly closed buds were 1.5 cm high and approximately 0.9 cm wide with pale pink petals.

20 Frozen tissue (1-3 g) was ground in liquid nitrogen with a mortar and pestle, placed in 25 mL pre-chilled Buffer A [0.2 M boric acid, 10 mM EDTA (sodium salt) (pH 7.6)] and homogenized briefly. The extract was mixed on a rotary shaker until it reached room temperature and an equal volume of phenol/chloroform (1:1 v/v), equilibrated with Buffer A, was added. After mixing for a further 10 minutes, the RNA preparation was centrifuged
25 at 10,000 x g for 10 minutes at 20°C. The upper aqueous phase was retained and the phenol interface re-extracted as above. The aqueous phases were pooled and adjusted to 0.1 M sodium acetate (pH 6.0), 2.5 volumes 95 % ethanol were added and the mixture was stored at -20°C overnight.

30 The preparation was centrifuged at 10,000 x g for 10 minutes at 4°C, the pellet dissolved

- 75 -

gently in 20 mL Buffer B [25 mM boric acid, 1.25 mM EDTA (sodium salt), 0.1 M NaCl (pH 7.6)] and 0.4 volumes 2-butoxyethanol (2BE) were added. This solution was incubated on ice for 30 minutes. It was then centrifuged at 10,000 x g for 10 minutes at 0°C and the supernatant was carefully collected. After addition of 1.0 volume of 2BE and incubation on ice for a further 30 minutes, the supernatant was again centrifuged at 10,000 x g for 10 minutes at 0°C. The resulting pellet was gently washed with Buffer A:2BE (1:1 v/v), then with 70% (v/v) ethanol, 0.1 M potassium acetate and finally with 95% ethanol. The pellet was air dried and dissolved in 1 mL diethyl pyrocarbonate (DEPC)-treated water. This was adjusted to 3 M lithium chloride, left on ice for 60 minutes and centrifuged at 10,000 x g for 10 minutes at 0°C. The pellet was washed twice with 3 M LiCl and then with 70% ethanol, 0.1 M potassium acetate.

The resulting RNA pellet was dissolved in 400 µL DEPC-treated water and extracted with an equal volume phenol/chloroform. The RNA mix was then centrifuged at 10,000 x g for 5 minutes at 20°C, the aqueous phase collected and made to 0.1 M sodium acetate, and a further 2.5 volumes of 95% ethanol were added. After 30 minutes incubation on ice, the mix was centrifuged at 13,000 rpm (5,000 x g) for 20 minutes at 20°C and the RNA pellet resuspended gently in 400 µL DEPC-treated water.

20 Poly (A)⁺ RNA was selected from the total RNA by Oligotex dT-30 (Takara, Japan) following the manufacturer's protocol. The cDNA was synthesized according to the method in Brugliera *et al.* (1994) and used to construct a non-directional petal cDNA library in the EcoRI site of λZAPII (Stratagene). The total number of recombinants obtained was 3.5 x 10⁵.

25

After transfecting XL1-Blue cells, the packaged cDNA mixture was plated at 50,000 pfu per 15 cm diameter plate. The plates were incubated at 37°C for 8 hours, and the phage were eluted in 100 mM NaCl, 8 mM MgSO₄, 50 mM Tris-HCl pH 8.0, 0.01% (w/v) gelatin (Phage Storage Buffer (PSB)) (Sambrook *et al.*, 1989). Chloroform was added and the phage stored at 4°C as an amplified library.

- 76 -

200,000 pfus of the amplified library were plated onto NZY plates (Sambrook *et al.*, 1989) at a density of 10,000 pfu per 15 cm plate after transfecting XL1-Blue MRF' cells, and incubated at 37°C for 8 hours. After incubation at 4°C overnight, duplicate lifts (labelled as group A and group B) were taken onto Colony/Plaque Screen™ filters (DuPont) and treated
5 as recommended by the manufacturer.

Screening of Kardinal cDNA library for a F3'H cDNA clone

Prior to hybridization, the duplicate plaque lifts were washed in prewashing solution (50 mM Tris-HCl pH7.5, 1 M NaCl, 1 mM EDTA, 0.1 % (w/v) sarcosine) at 65°C for 30 minutes;
10 stripped in 0.4 M sodium hydroxide at 65°C for 30 minutes; then washed in a solution of 0.2 M Tris-HCl pH 8.0, 0.1 x SSC, 0.1 % (w/v) SDS at 65°C for 30 minutes and finally rinsed in 2 x SSC, 1.0 % (w/v) SDS.

The group A filters of the duplicate lifts from the Kardinal cDNA library were screened with
15 ³²P-labelled fragments of an NcoI fragment from pCGP1805 containing the petunia Htl (OGR-38) cDNA clone, while the group B filters were screened with ³²P-labelled fragments of EcoRI/SspI fragment from pCGP246 containing the snapdragon F3'H clone.

Hybridization conditions included a prehybridization step in 10 % (v/v) formamide, 1 M
20 NaCl, 10 % (w/v) dextran sulphate, 1 % (w/v) SDS at 42°C for at least 1 hour. The ³²P-labelled fragment (2x10⁶cpm/mL) was then added to the hybridization solution and hybridization was continued at 42°C for a further 16 hours. The filters were then washed at 42°C in 2 x SSC, 1 % (w/v) SDS for 2 hours followed by 1 x SSC, 1 % (w/v) SDS for 1 hour and finally in 0.2 x SSC/1 % (w/v) SDS for 2 hours. The filters were exposed to
25 Kodak XAR film with an intensifying screen at -70°C for 16 hours.

Four strongly-hybridizing plaques (R1, R2, R3, R4) were picked into PSB and rescreened to isolate pure plaques. The plasmids contained in the λZAP bacteriophage vector were rescued and digested with EcoRI to release the cDNA inserts. Clone R1 contained a 1.0 kb
30 insert while clones R2, R3 and R4 contained inserts of approximately 1.3 kb each. Sequence

- 77 -

data were generated from the 3' and 5' ends of the R4 cDNA insert.

The rose R4 putative F3'H sequence was compared with that of the petunia OGR-38 F3'H sequence. At the nucleotide level, the R4 cDNA clone showed 63.2 % and 62.1 % similarity over 389 nucleotides at the 5' end and 330 nucleotides at the 3' end, respectively. At the amino acid level, the R4 clone showed 65.4 % and 73.9 % similarity over 130 amino acids at the 5' end and 69 amino acids at the 3' end, respectively. Based on the high sequence similarity of the Rose R4 cDNA clone to that of the petunia F3'H cDNA clone (OGR-38), a corresponding "full-length" cDNA clone was isolated, as described in Example 25, below.

10

EXAMPLE 25- Isolation of a full-length rose F3'H cDNA

In order to isolate a "full-length" F3'H cDNA clone from Rose, the *Rosa hybrida* cv Kardinal petal cDNA library described in Example 24 was screened with ³²P-labelled fragments of the rose R4 cDNA clone, described above.

15

A total of 1.9×10^6 pfus of the amplified library were plated onto NZY plates at a density of 100,000 pfus per 15 cm diameter plate after transfecting XL1-Blue MRF' cells, and incubated at 37°C for 8 hours. After incubation at 4°C overnight, duplicate lifts were taken onto Colony/Plaque Screen™ filters (DuPont) and treated as recommended by the manufacturer.

20

Screening of Kardinal cDNA library for full-length F3'H cDNA clones

Prior to hybridization, the duplicate plaque lifts were treated as described in Example 24.

25 The duplicate lifts from the Kardinal cDNA library were screened with ³²P-labelled fragments of an EcoRI fragment from the rose R4 cDNA clone.

Hybridization conditions included a prehybridization step in 50 % (v/v) formamide, 1 M NaCl, 10 % (w/v) dextran sulphate, 1 % (w/v) SDS at 42°C for at least 1 hour. The ³²P-labelled fragment of the rose R4 cDNA clone (1×10^6 cpm/mL) was then added to the

30

- 78 -

hybridization solution and hybridization was continued at 42°C for a further 16 hours. The filters were then washed in 2 x SSC, 1% (w/v) SDS at 42°C for 2 x 1 hour and exposed to Kodak XAR film with an intensifying screen at -70°C for 16 hours.

- 5 Seventy-three strongly-hybridizing plaques (1-73) were picked into 1mL of PSB and stored at 4°C overnight. 100µL of each was then aliquoted into a microtitre tray as an ordered array.

XL1-Blue MRF' cells were added to 10mL of molten NZY top agar, poured onto NZY
10 plates (15cm diameter) and allowed to set. A replica plating device was used to transfer the 73 phage isolates in an ordered array onto the NZY plate previously inoculated with the XL1-Blue MRF' cells. After incubation at 37°C for 6 hours followed by 4°C overnight, triplicate lifts (arrays 1, 2 and 3) were taken onto Colony/Plaque Screen™ filters (DuPont) and treated as recommended by the manufacturer.

15

Prior to hybridization, the duplicate plaque lifts were treated as described in Example 24.

The 3 arrays were screened with ³²P-labelled fragments of a) an EcoRI/SalI fragment covering the 5' end of the rose R4 cDNA clone, b) an EcoRI/ClaI fragment covering the 5'
20 end of the rose R4 cDNA clone or c) an EcoRI fragment of the entire rose R4 cDNA clone using the hybridisation and washing conditions described above, except that the final wash was in 0.1 x SSC, 0.1% (w/v) SDS at 65°C for 30 minutes. The filters were exposed to Kodak XAR film with an intensifying screen at -70°C for 16 hours.

- 25 All 73 plaques hybridised with the full R4 cDNA clone (EcoRI fragment) whilst a total of only 17 hybridised with the 5' end of the R4 cDNA clone (either EcoRI/SalI or the EcoRI/ClaI fragments). The 17 phage isolates were rescreened as described above to isolate purified plaques. Pure plaques were obtained from 9 out of the 17 (2, 4, 26, 27, 34, 38, 43, 44, 56). The plasmids contained in the λZAP bacteriophage vector were rescued and the
30 sizes of the cDNA inserts were determined using an EcoRI digestion. The cDNA inserts

- 79 -

ranged from 0.9kb to 1.9kb. Of the nine, only #34 (named pCGP2158) and #38 (named pCGP2159) contained cDNA inserts of approximately 1.9kb. Sequence data were generated from the 3' and 5' ends of the cDNA inserts and showed that clones #34 and #38 represented the same gene.

5

The complete sequence of the rose cDNA clone (#34) contained in the plasmid pCGP2158 was determined by compilation of sequence from different pUC18 subclones obtained using standard procedures for the generation of randomly-overlapping clones (Sambrook *et al.*, 1989). The sequence (SEQ ID NO:14) contained an open reading frame of 1696 bases
10 which encodes a putative polypeptide of 520 amino acids (SEQ ID NO:15).

The nucleotide and predicted amino acid sequences of the rose F3'H #34 cDNA clone (SEQ ID NO:14 and SEQ ID NO:15) were compared with those of the petunia OGR-38 F3'H cDNA clone (SEQ ID NO:1 and SEQ ID NO:2) and the snapdragon sdF3'H clone (SEQ ID
15 NO:3 and SEQ ID NO:4). The rose F3'H #34 cDNA clone showed 64.7% similarity, over 1651 nucleotides, and 72.7% similarity, over 509 amino acids, to that of the petunia OGR-38 cDNA clone, and 67.2% similarity, over 1507 nucleotides, and 68.9 similarity, over 502 amino acids, to that of the snapdragon sdF3'H clone.

20 An alignment of the petunia, carnation, snapdragon, arabidopsis, rose, chrysanthemum and torenia sequences, all of which are disclosed in this specification, and various summaries of comparisons of sequence similarities among the nucleotide and corresponding amino acid sequences, can be found in Table 7 and in Tables 8, 9, 10, 11 and 12, respectively. These Tables are in Example 34, at the end of the specification.

25

**EXAMPLE 26- Stable expression of the rose F3'H cDNA clone (#34) in petunia petals-
Complementation of a *ht1/ht1* petunia cultivar**

30 **Preparation of pCGP2166**

- 80 -

Plasmid pCGP2166 (Figure 16) was constructed by cloning the cDNA insert from pCGP2158 in a "sense" orientation behind the Mac promoter (Comai *et al.*, 1990) of pCGP293 (Brugliera *et al.*, 1994). The plasmid pCGP2158 was digested with EcoRI to release the cDNA insert. The overhanging 5' ends were filled in using DNA polymerase 5 (Klenow fragment) (Sambrook *et al.*, 1989). The cDNA fragment was isolated and ligated with filled in BamHI ends of the pCGP293 binary vector. Correct insertion of the fragment in pCGP2166 was established by restriction enzyme analysis of DNA isolated from gentamycin-resistant transformants.

10 The binary vector pCGP2166 was introduced into *A. tumefaciens* strain AGL0 cells, as described in Example 9. The pCGP2166/AGL0 cells were then used to transform Skr4 x SW63 petunia plants (also described in Example 9), to test for stable expression and activity of the enzyme encoded by the gene corresponding to the rose #34 cDNA clone.

15 **EXAMPLE 27- Transgenic plant phenotype analysis**

pCGP2166 in Skr4 x SW63

The expression of the introduced rose F3'H cDNA in the Skr4 x SW63 hybrid had a marked effect on flower colour. The stamen tissue of the non-transgenic control is white, whereas the same tissue in most of the transgenic plants was pink. In addition, expression of the rose 20 F3'H cDNA in the Skr4 x SW63 hybrid conferred a dark pink hue (RHSCC# 64C and 74C) to the corolla, which is normally pale lilac (RHSCC# 75C). The colour codes are taken from the Royal Horticultural Society's Colour Chart (RHSCC). They provide an alternative means by which to describe the colour phenotypes observed. The designated numbers, however, should be taken only as a guide to the perceived colours and should not be regarded 25 as limiting the possible colours which may be obtained.

Acid-hydrolysed floral extracts (see Example 11) were run in a Forestal solvent system (HOAc:water:HCl; 30: 10: 3) (Markham, 1982). The 3' hydroxylated flavonoids, peonidin and quercetin, were readily detected in the petal limbs of the transgenic plants. Only 30 kaempferol and a small amount of malvidin were detected in the non-transgenic Skr4 x SW63

control.

The accumulation of the 3'-hydroxylated anthocyanidin, peonidin and the flavonol, quercetin, in the petals of the transgenic Skr4 x SW63/pCGP2166 plants correlated with the pink and dark pink colours observed in the petals of the same plants.

Preparation of pCGP2169

The binary construct pCGP2169 (Figure 17) was prepared by cloning the cDNA insert from pCGP2158 in a "sense" orientation between the CaMV35S promoter (Franck *et al.*, 1980; Guilley *et al.*, 1982) and *ocs* terminator (De Greve *et al.*, 1982). The plasmid pCGP1634 contained a CaMV35S promoter, β -glucuronidase (GUS) reporter gene encoded by the *E. coli uidA* locus (Jefferson *et al.*, 1987) and *ocs* terminator region in a pUC19 vector. The plasmid pCGP2158 was digested with NcoI/XbaI to release the cDNA insert. The plasmid pCGP1634 was also digested with NcoI/XbaI to release the backbone vector containing the CaMV35S promoter and the *ocs* terminator. The fragments were isolated and ligated together to produce pCGP2167. The plasmid pCGP2167 was subsequently digested with PvuII to release the expression cassette containing the CaMV35S promoter, the rose F3'H cDNA clone and the *ocs* terminator. This expression cassette fragment was isolated and ligated with SmaI ends of pWTT2132 binary vector (DNA Plant Technology Corporation; Oakland, California) to produce pCGP2169 (Figure 17).

The binary vector pCGP2169 was introduced into *A. tumefaciens* strain AGL0 cells, as described in Example 9. The pCGP2169/AGL0 cells are used to transform rose plants, to reduce the amount of 3'-hydroxylated flavonoids.

EXAMPLE 28- Isolation of a putative F3'H cDNA clone from chrysanthemum

In order to isolate a chrysanthemum F3'H cDNA clone, a chrysanthemum cv. Red Minstral petal cDNA library was screened with ³²P-labelled fragments of the petunia Ht1 cDNA clone (OGR-38), contained in pCGP1805.

Construction of a petal cDNA library from chrysanthemum cv. Red Minstral

Total RNA was prepared from the petals (stages 3 to 5) of chrysanthemum cv. Red Minstral using Trizol™ reagent (Life Technologies) (Chomczynski and Sacchi, 1987) according to the manufacturer's recommendations. Poly(A)⁺ RNA was enriched from the total RNA, using
5 a mRNA isolation kit (Pharmacia) which relies on oligo-(dT) affinity spun-column chromatography .

A Superscript™ cDNA synthesis kit (Life Technologies) was used to construct a petal cDNA library in ZipLox using 5 µg of poly(A)⁺ RNA isolated from stages 3 to 5 of Red Minstral
10 as template.

30,000 pfus of the library were plated onto LB plates (Sambrook *et al.*, 1989) at a density of 3,000 pfus per 15 cm plate after transfecting Y1090r-, and incubated at 37°C for 16 hours. After incubation at 4°C for one hour, duplicate lifts were taken onto Hybond
15 N+™ filters (Amersham) and treated as recommended by the manufacturer.

Screening of the Red Minstral cDNA Library

The duplicate lifts from the Red Minstral petal cDNA library were screened with ³²P-labelled fragments of the 1.8 kb Asp718/BamHI insert from pCGP1805.

20

Hybridization conditions included a prehybridization step in 1mM EDTA (pH8.0), 0.5M Na₂HPO₄ (pH7.2), 7% (w/v) SDS (Church and Gilbert, 1984) at 65°C for at least 1 hour. The ³²P-labelled fragments (1x10⁶cpm/mL) were then added to the hybridization solution and hybridization was continued at 65°C for a further 16 hours. The filters were
25 then washed in 2 x SSC, 0.1% (w/v) SDS at 65°C for 2 x 1 hour and exposed to Kodak BioMax™ film with an intensifying screen at -70°C for 48 hours.

Eight strongly-hybridizing plaques were picked into PSB (Sambrook *et al.*, 1989). Of these, 2 (RM6i and RM6ii) were rescreened to isolate purified plaques, using the hybridization
30 conditions as described for the initial screening of the cDNA library. The plasmids

contained in the λ ZipLox bacteriophage vector were rescued according to the manufacturer's protocol and sequence data was generated from the 3' and 5' ends of the cDNA inserts. The partial sequences of the RM6i and RM6ii cDNA inserts were compared with the complete sequence of the petunia OGR-38 F3'H cDNA clone. The RM6i cDNA clone showed
5 relatively high sequence similarity with that of the petunia OGR-38 cDNA clone, and was further characterised.

The RM6i cDNA insert contained in pCHRM1 was released upon digestion with EcoRI and was approximately 1.68 kb. The complete sequence of RM6i cDNA clone (SEQ ID NO:16)
10 contained in the plasmid pCHRM1 was determined by compilation of sequence from subclones of the RM6i cDNA insert.

The nucleotide and predicted amino acid sequences of the chrysanthemum RM6i cDNA insert (SEQ ID NO:16 and SEQ ID NO:17) were compared with those of the petunia OGR-
15 38 F3'H cDNA clone (SEQ ID NO:1 and SEQ ID NO:2). The sequence of the chrysanthemum RM6i cDNA insert showed 68.5% similarity, over 1532 nucleotides, and 73.6 % similarity, over 511 amino acids, to that of the petunia OGR-38 F3'H cDNA clone.

An alignment of the petunia, carnation, snapdragon, arabidopsis, rose, chrysanthemum and
20 torenia sequences, all of which are disclosed in this specification, and various summaries of comparisons of sequence similarities among the nucleotide and corresponding amino acid sequences, can be found in Table 7 and in Tables 8, 9, 10, 11 and 12, respectively. These Tables are in Example 34, at the end of the specification.

25 Construction of pLN85 (antisense binary)

A plasmid designated pLN84 was constructed by cloning the RM6i cDNA insert from pCHRM1 in the "antisense" orientation behind the complete CaMV35S promoter contained in pART7 (Gleave 1992). The plasmid pCHRM1 was digested with NotI to release the cDNA insert. The RM6i cDNA fragment was blunt-ended using T4 DNA polymerase
30 (Sambrook et al., 1989) and purified, following agarose gel electrophoresis and GELase

- 84 -

(Epicentre Technologies). The purified fragment was ligated with SmaI ends of the pART7 shuttle vector to produce pLN84. The plasmid pLN84 was subsequently digested with NotI to release the expression cassette containing CaMV35S: RM6i cDNA: ocs. The expression cassette was isolated as a single fragment and ligated with NotI ends of the pART27 binary
5 vector (Gleave, 1992) to produce pLN85 (Figure 18). Correct insertion of the fragment was established by restriction enzyme analysis of DNA isolated from streptomycin-resistant *E. coli* transformants.

The binary vector pLN85 is introduced into chrysanthemum plants via *Agrobacterium*-
10 mediated transformation, as described in Ledger *et al*, 1991), to reduce the amount of 3'-hydroxylated flavonoids.

EXAMPLE 29- Isolation of a putative F3'H cDNA clone from *Torenia fournieri*

15 In order to isolate a torenia F3'H cDNA clone, the petunia Ht1-linked F3'H cDNA clone (OGR-38), contained in pCGP1805, was used to screen a *Torenia fournieri* cv. Summer Wave petal cDNA library, under low stringency conditions.

Construction of *Torenia fournieri* cv. Summer Wave petal cDNA library

20 A directional petal cDNA library was prepared from Summer Wave flowers, essentially as described in Example 4.

Screening of Summer Wave petal cDNA library

Lifts of a total of 200,000 of the amplified Summer Wave petal cDNA library were screened
25 with DIG-labelled fragments of the 1.8 kb OGR-38 cDNA insert from pCGP1805. A DIG DNA labelling and detection kit from Boehringer-Mannheim was used according to the manufacturer's recommendations.

Hybridizations were carried out in 30% (v/v) formamide, 5 x SSC, 1% (w/v) SDS at 37°C
30 for 16 hours. The filters were then washed in 5 x SSC, 1% (w/v) SDS at 65°C for 1 hour.

The signals were visualized following the protocol of the DIG DNA labelling and detection kit.

Twelve strongly-hybridizing plaques were picked into PSB and rescreened to isolate pure
5 plaques. The plasmids contained in the λ ZAPII bacteriophage vector were rescued and
digested with EcoRI/XhoI to release the cDNA inserts. Most of the twelve clones contained
cDNA inserts of approximately 1.8 kb. One clone, THT52, contained the longest 5' non-
coding-region sequence. The complete sequence of the torenia cDNA clone (THT52),
contained in the plasmid pTHT52, was determined by compilation of sequence from different
10 pUC18 subclones obtained using standard procedures for the generation of randomly-
overlapping clones (Sambrook *et al.*, 1989). The sequence (SEQ ID NO:18) contained an
open reading frame of 1524 bases which encodes a putative polypeptide of 508 amino acids
(SEQ ID NO:19).

15 The nucleotide and predicted amino acid sequences of the torenia THT52 cDNA clone (SEQ
ID NO:18 and SEQ ID NO:19) were compared with those of the petunia OGR-38 F3'H
cDNA clone (SEQ ID NO:1 and SEQ ID NO:2). The torenia THT52 cDNA clone showed
63.6% similarity, over 1694 nucleotides, and 67.4% similarity, over 515 amino acids, to that
of the petunia OGR-38 cDNA clone.

20

An alignment of the petunia, carnation, snapdragon, arabidopsis, rose, chrysanthemum and
torenia sequences, all of which are disclosed in this specification, and various summaries of
comparisons of sequence similarities among the nucleotide and corresponding amino acid
sequences, can be found in Table 7 and in Tables 8, 9, 10, 11 and 12, respectively. These
25 Tables are in Example 34, at the end of the specification.

**EXAMPLE 30- The F3'H assay of the torenia THT cDNA clone expressed in yeast
Construction of pYTHT6**

30 The plasmid pYTHT6 (Figure 19) was constructed by cloning the cDNA insert from pTHT6

- 86 -

in a "sense" orientation behind the yeast glyceraldehyde-3-phosphate dehydrogenase promoter of pYE22m (Tanaka *et al.*, 1988). The plasmid pTHT6 contained the THT6 cDNA clone. THT6 is identical to THT52, except that its 5' non-coding region is 75 bp shorter.

- 5 The 1.7kb THT6 cDNA insert was released from the plasmid pTHT6 upon digestion with EcoRI/XhoI. The THT6 cDNA fragment was isolated, purified and ligated with EcoRI/SalI ends of pYE22m to produce pYTHT6.

Yeast transformation, preparation of yeast extracts and the F3'H assay are described in
10 Example 6.

F3'H activity was detected in extracts of G1315/pYTHT6, but not in extracts of non-transgenic yeast. From this it was concluded that the THT6 cDNA insert contained in pYTHT6, encoded a F3'H.

15

EXAMPLE 31- Isolation of a putative F3'H cDNA clone from *Pharbitis nil* (Japanese morning glory)

In order to isolate a morning glory F3'H cDNA clone, the petunia Ht1-linked F3'H cDNA
20 clone (OGR-38), contained in pCGP1805, was used to screen a Japanese morning glory petal cDNA library, under low stringency conditions.

Construction of Japanese morning glory petal cDNA library

The petal cDNA library from young petals of *Pharbitis nil* (Japanese morning glory) was
25 obtained from Dr Iida (National Institute of Basic Biology, Japan).

Screening of Japanese morning glory petal cDNA library

Lifts of a total of 200,000 of the amplified Japanese morning glory petal cDNA library were screened with DIG-labelled fragments of the 1.8 kb OGR-38 cDNA insert from pCGP1805.
30 A DIG DNA labelling and detection kit from Boehringer-Mannheim was used according to

the manufacturer's recommendations.

Hybridizations were carried out in 30% (v/v) formamide, 5 x SSC, 1% (w/v) SDS at 37°C for 16 hours. The filters were then washed in 5 x SSC, 1% (w/v) SDS at 65°C for 1 hour.

5 The signals were visualized following the protocol of the DIG DNA labelling and detection kit.

Twenty strongly-hybridizing plaques were picked into PSB and rescreened to isolate pure plaques. The plasmids contained in the λ ZAPII bacteriophage vector were rescued and
10 digested with EcoRI/XhoI to release the cDNA inserts. One clone (MHT85) contained a 1.8kb insert. The complete sequence of the Japanese morning glory cDNA clone (MHT85) (SEQ ID NO:20), contained in the plasmid pMHT85, was determined by compilation of sequence from different pUC18 subclones obtained using standard procedures for the generation of randomly-overlapping clones (Sambrook *et al.*, 1989). The MHT85 sequence
15 appears to be 5 bases short of "full-length".

The nucleotide and predicted amino acid sequences of the Japanese morning glory MHT85 cDNA clone (SEQ ID NO:20 and SEQ ID NO:21) were compared with those of the petunia OGR-38 F3'H cDNA clone (SEQ ID NO:1 and SEQ ID NO:2). The Japanese morning
20 glory MHT85 cDNA clone showed 69.6% similarity, over 869 nucleotides, and 74.8% similarity, over 515 amino acids, to that of the petunia OGR-38 cDNA clone.

An alignment of the petunia, carnation, snapdragon, arabidopsis, rose, chrysanthemum and torenia sequences, all of which are disclosed in this specification, and various summaries of
25 comparisons of sequence similarities among the nucleotide and corresponding amino acid sequences, can be found in Table 7 and in Tables 8, 9, 10, 11 and 12, respectively. These Tables are in Example 34, at the end of the specification.

30 **EXAMPLE 32- Isolation of a putative F3'H cDNA clone from *Gentiana triflora***

In order to isolate a gentian F3'H cDNA clone, the petunia Ht1-linked F3'H cDNA clone (OGR-38), contained in pCGP1805, was used to screen a *Gentiana triflora* Pall. var *japonica* Hara petal cDNA library, under low stringency conditions.

5 Construction of gentian petal cDNA library

A petal cDNA library was prepared from *Gentiana triflora* Pall. var *japonica* Hara flowers, as described by Tanaka *et al.*, 1996.

Screening of gentian petal cDNA library

- 10 Lifts of a total of 200,000 of the amplified gentian petal cDNA library were screened with DIG-labelled fragments of the 1.8 kb OGR-38 cDNA insert from pCGP1805. A DIG DNA labelling and detection kit from Boehringer-Mannheim was used according to the manufacturer's recommendations.
- 15 Hybridizations were carried out in 30% (v/v) formamide, 5 x SSC, 1% (w/v) SDS at 37°C for 16 hours. The filters were then washed in 5 x SSC, 1% (w/v) SDS at 65°C for 1 hour. The signals were visualized following the protocol of the DIG DNA labelling and detection kit.
- 20 Fifteen strongly-hybridizing plaques were picked into PSB and rescreened to isolate pure plaques. The plasmids contained in the λ ZAPII bacteriophage vector were rescued and digested with EcoRI/XhoI to release the cDNA inserts. One clone (GHT13) contained a 1.8kb insert. The sequence of the partial gentian cDNA clone (GHT13) (SEQ ID NO:22), contained in the plasmid pGHT13, was determined by compilation of sequence from different
25 pUC18 subclones obtained using standard procedures for the generation of randomly-overlapping clones (Sambrook *et al.*, 1989).

The nucleotide and predicted amino acid sequences of the gentian GHT13 cDNA clone (SEQ ID NO:22 and SEQ ID NO:23) were compared with those of the petunia OGR-38 F3'H
30 cDNA clone. The gentian GHT13 cDNA clone showed 68.3% similarity, over 1519

- 89 -

nucleotides, and 71.8% similarity, over 475 amino acids, to that of the petunia OGR-38 cDNA clone.

An alignment of the petunia, carnation, snapdragon, arabidopsis, rose, chrysanthemum and torenia sequences, all of which are disclosed in this specification, and various summaries of comparisons of sequence similarities among the nucleotide and corresponding amino acid sequences, can be found in Table 7 and in Tables 8, 9, 10, 11 and 12, respectively. These Tables are in Example 34, at the end of the specification.

10

EXAMPLE 33- Isolation of putative F3'H cDNA clone from lisianthus

In order to isolate a lisianthus F3'H cDNA clone, the petunia Ht1-linked F3'H cDNA clone (OGR-38), contained in pCGP1805, was used to screen a lisianthus petal cDNA library, under low stringency conditions.

15

Construction and screening of lisianthus petal cDNA library

10,000 pfus of a lisianthus petal cDNA library described by Davies *et al.* (1993) and Markham and Offman (1993) were plated onto LB plates (Sambrook *et al.*, 1989) at a density of 3,000 pfus per 15 cm plate after transfecting Y1090r-, and incubated at 37°C for 16 hours. After incubation at 4°C for one hour, duplicate lifts were taken onto Hybond N+™ filters (Amersham) and treated as recommended by the manufacturer.

The duplicate lifts from the lisianthus line #54 petal cDNA library were screened with ³²P-labelled fragments of the 1.8 kb Asp718/BamHI insert from pCGP1805.

25

Hybridization conditions included a prehybridization step in 1mM EDTA (pH8.0), 0.5MNa₂HPO₄ (pH7.2), 7% (w/v) SDS (Church and Gilbert, 1984) at 55°C for at least 1 hour. The ³²P-labelled fragments (1x10⁶cpm/mL) were then added to the hybridization solution and hybridization was continued at 55°C for a further 16 hours. The filters were then washed in 2 x SSC, 0.1% (w/v) SDS at 55°C for 2 x 15 minutes, and exposed to Kodak

- 90 -

BioMaxTM film with an intensifying screen at -70°C for 18 hours.

Twelve strongly-hybridizing plaques were picked into PSB (Sambrook *et al.*, 1989) and rescreened to isolate purified plaques, using the hybridization conditions as described for the
5 initial screening of the cDNA library. Sequence data were generated from the 3' and 5' ends of the cDNA inserts of four clones.

Based on sequence comparisons, pL3-6 showed similarity with the petunia OGR-38 F3'H cDNA clone and was further characterised.

10

The 2.2 kb cDNA insert, contained in pL3-6, was subsequently found to contain 3 truncated cDNA clones, the longest (L3-6) having high sequence similarity to the petunia OGR-38 cDNA sequence. The sequence of this L3-6 partial cDNA clone contained in the plasmid pL3-6 was determined by compilation of sequence from subclones of the L3-6 cDNA insert
15 (SEQ ID NO:24).

The nucleotide and predicted amino acid sequences of the lisianthus L3-6 cDNA clone (SEQ ID NO:24 and SEQ ID NO:25) were compared with those of the petunia OGR-38 F3'H cDNA clone (SEQ ID NO:1 and SEQ ID NO:2). The sequence of the lisianthus L3-6 cDNA
20 clone showed 71.4% similarity, over 1087 nucleotides, and 74.6% similarity, over 362 amino acids, to that of the petunia OGR-38 F3'H cDNA clone .

An alignment of the petunia, carnation, snapdragon, arabidopsis, rose, chrysanthemum and
25 torenia sequences, all of which are disclosed in this specification, and various summaries of comparisons of sequence similarities among the nucleotide and corresponding amino acid sequences, can be found in Table 7 and in Tables 8, 9, 10, 11 and 12, respectively. These Tables are in Example 34, at the end of the specification.

30 Further investigation of the remaining clones isolated from the screening of the lisianthus

- 91 -

library identified another putative F3'H cDNA clone (L3-10), contained in the plasmid pL3-10. The L3-10 cDNA insert is approximately 1.8kb and appears to represent a "full-length" clone.

5 EXAMPLE 34-Alignments and comparisons among nucleotide and amino acid sequences disclosed herein

Multiple sequence alignments were performed using the ClustalW program as described in Example 3. Table 7 (below) provides a multiple sequence alignment of the predicted amino acid sequences of petunia OGR-38 (A); carnation (B); snapdragon (C); arabidopsis Tt7
10 coding region (D); rose (E) chrysanthemum (F); torenia (G); morning glory (H); gentian (partial sequence) (I); lisianthus (partial sequence) (J) and the petunia 651 cDNA (K). Conserved amino acids are shown in bolded capital letters and are boxed and shaded. Similar amino acids are shown in capital letters and are only lightly shaded, and dissimilar amino acids are shown in lower case letters.

15

Nucleotide and amino acid sequences of the F3'H cDNA clones from the above mentioned species and the coding region of the genomic clone from arabidopsis were compared using the LFASTA program, as described in Example 3. Summaries of similarity comparisons are presented in Tables 8 to 12, below.

20

TABLE 7

i
ii
iii
iv
v

-93-

i

A	/									m e i l s l	I	l y t	V	i f s f l	L	q f i	L	21
B	/												M h n l y	Y L	i t t	V	11	
C	/	m q h q y y s l i t m d d i s								I	t s l	L v p c t	F I	l g f	L		30	
D	/									m a t l f	L t i l	L a t v l	F L	i l r	I		20	
E	/									m f l i	V v i t f l f a v	F L	f r l	L		19		
F	/									m t i l a f	V f y a	L i l g s v	L y v f	L		21		
G	/									m s p l a l m i	I s t l	L g f l l	Y h s l r	L		23		
H	/									s l t l	I f c t	L v f a i	F L	y f l	I	19		
I	/																0	
J	/																0	
K	/									m d y v n	I l l g	L f f t w	F L	v n g	L	20		

A	22	r - s f f r k	R	y p l p	L P P G P K P W P I I G N L v	H L G	50
B	12	f r g - - -	-	h q k p	L P P G P R P W P I V G N L P H M G		36
C	31	l l y s f l n	K	k v k p	L P P G P K P W P I V G N L P H L G		60
D	21	f s h r r n r	s h n n r		L P P G P n P W P I I G N L P H M G		50
E	20	f s g k s q r	- h s l p		L P P G P K P W P V V G N L P H L G		48
F	22	n l s - - -	s	R k s a r	L P P G P t P W P I V G N L P H L G		48
G	24	l l f s g q g	R	- - r l	L P P G P R P W P L V G N L P H L G		51
H	20	l r - - v k q	R	y p l p	L P P G P K P W P V L G N L P H L G		47
I	1					P I L G N I P H L G	10
J	1						0
K	21	m s l r - r r	K	i s k k	L P P G P f P l	P I I G N L h l L G	49

A	51	p	K	P	H	Q	S	t	A	A	M	A	Q	t	Y	G	P	L	M	y	L	K	M	G	F	V	D	V	V	V	A	80
B	37	q	a	P	H	Q	g	L	A	A	L	A	Q	k	Y	G	P	L	L	y	M	R	L	G	Y	V	D	V	V	V	A	66
C	61	p	K	P	H	Q	S	M	A	A	L	A	R	v	h	G	P	L	I	H	L	K	M	G	F	V	h	V	V	V	A	90
D	51	t	K	P	H	R	T	L	S	A	M	v	t	t	Y	G	P	I	L	H	L	R	L	G	F	V	D	V	V	V	A	80
E	49	p	f	P	H	H	S	I	A	e	L	A	K	k	h	G	P	L	M	H	L	R	L	G	Y	V	D	V	V	V	A	78
F	49	p	i	P	H	H	A	L	A	A	L	A	K	k	Y	G	P	L	M	H	L	R	L	G	c	V	D	V	V	V	A	78
G	52	p	K	P	H	a	S	M	A	e	L	A	R	a	Y	G	P	L	M	H	L	K	M	G	F	V	h	V	V	V	A	81
H	48	k	K	P	H	Q	S	I	A	A	M	A	e	r	Y	G	P	L	M	H	L	R	L	G	F	V	D	V	V	V	A	77
I	11	s	K	P	H	Q	T	L	A	e	M	A	K	t	Y	G	P	L	M	H	L	K	f	G	l	k	D	a	V	V	A	40
J	1																														0	
K	50	n	H	P	H	K	S	L	A	q	L	A	K	i	h	G	P	I	M	N	L	K	L	G	q	L	n	t	V	V	i	79

A	81	A	S	A	S	V	A	a	Q	F	L	K	t	H	D	A	N	F	S	S	R	P	P	N	S	G	A	e	H	M	A	110
B	67	A	S	A	S	V	A	t	Q	F	L	K	t	H	D	l	N	F	S	S	R	P	P	N	S	G	A	K	H	I	A	96
C	91	S	S	A	S	V	A	e	k	F	L	K	v	H	D	A	N	F	S	S	R	P	P	N	S	G	A	K	H	V	A	120
D	81	A	S	k	S	V	A	e	Q	F	L	K	i	H	D	A	N	F	A	S	R	P	P	N	S	G	A	K	H	M	A	110
E	79	A	S	A	S	V	A	a	Q	F	L	K	t	H	D	A	N	F	S	S	R	P	P	N	S	G	A	K	H	L	A	108
F	79	A	S	A	S	V	A	a	Q	F	L	K	v	H	D	A	N	F	A	S	R	P	P	N	S	G	A	K	H	V	A	108
G	82	S	S	A	S	a	A	e	Q	c	L	R	v	H	D	A	N	F	l	S	R	P	P	N	S	G	A	K	H	V	A	111
H	78	A	S	A	A	V	A	a	Q	F	L	K	v	H	D	S	N	F	S	n	R	P	P	N	S	G	A	e	H	I	A	107
I	41	S	S	A	S	V	A	e	Q	F	L	K	k	H	D	v	N	F	S	n	R	P	P	N	S	G	A	K	H	I	A	70
J	1																														0	
K	80	S	S	S	v	V	A	r	E	v	L	Q	k	Q	D	l	T	F	S	n	R	f	v	p	d	v	v	H	v	r	n	109

-94-
ii

A	111	Y	N	Y	Q	D	L	V	F	A	P	Y	G	P	R	W	R	M	L	R	K	I	C	S	V	H	L	F	S	T	K	140
B	97	Y	N	Y	Q	D	L	V	F	A	P	Y	G	P	K	W	R	M	L	R	K	I	C	S	L	H	M	F	S	S	K	126
C	121	Y	N	Y	Q	D	L	V	F	A	P	Y	G	P	R	W	R	M	L	R	K	I	C	A	L	H	L	F	S	A	K	150
D	111	Y	N	Y	Q	D	L	V	F	A	P	Y	G	h	R	W	R	L	L	R	K	I	S	S	V	H	L	F	S	A	K	140
E	109	Y	N	Y	Q	D	L	V	F	r	P	Y	G	P	R	W	R	M	f	R	K	I	S	S	V	H	L	F	S	g	K	138
F	109	Y	N	Y	Q	D	L	V	F	A	P	Y	G	P	R	W	R	L	L	R	K	I	C	S	V	H	L	F	S	A	K	138
G	112	Y	N	Y	E	D	L	V	F	r	P	Y	G	P	K	W	R	L	L	R	K	I	C	A	q	H	I	F	S	v	K	141
H	108	Y	N	Y	Q	D	L	V	F	A	P	Y	G	P	R	W	R	M	L	R	K	I	T	S	V	H	L	F	S	A	K	137
I	71	Y	N	Y	Q	D	L	V	F	A	P	Y	G	P	R	W	R	L	L	R	K	I	C	S	V	H	L	F	S	S	K	100
J	1																														0	
K	110	h	s	d	f	s	V	V	w	l	P	v	N	s	R	W	K	t	L	R	K	I	m	n	s	s	I	F	S	g	n	139

A	141	A	L	D	D	F	R	H	V	R	Q	D	-	-	-	E	V	k	t	L	T	R	A	L	A	s	A	G	q	k	P	167	
B	127	A	L	D	D	F	R	l	V	R	Q	E	-	-	-	E	V	S	I	L	v	n	A	I	A	k	A	G	t	k	P	153	
C	151	A	L	n	D	D	F	t	H	V	R	Q	D	-	-	-	E	V	g	I	L	T	R	v	L	A	d	A	G	e	t	P	177
D	141	A	L	E	D	D	F	K	H	V	R	Q	E	-	-	-	E	V	g	t	L	T	R	e	L	v	r	v	G	t	k	P	167
E	139	A	L	D	D	l	K	H	V	R	Q	E	-	-	-	E	V	S	V	L	A	H	A	L	A	n	S	G	s	k	v	165	
F	139	A	L	D	D	F	R	H	V	R	Q	E	-	-	-	E	V	A	V	L	T	R	v	L	l	s	A	G	n	s	P	165	
G	142	A	M	D	D	F	R	R	V	R	E	E	-	-	-	E	V	A	I	L	S	R	A	L	A	-	G	k	r	a	166		
H	138	A	L	D	D	F	c	H	V	R	Q	E	-	-	-	E	V	A	t	L	T	R	S	L	A	s	A	G	k	t	P	164	
I	101	A	L	D	D	F	c	H	V	R	h	E	-	-	-	E	I	C	I	L	i	R	A	I	A	s	g	G	h	a	P	127	
J	1															r	I	L	T	R	S	I	A	s	A	G	e	n	P	14			
K	140	k	L	D	g	n	Q	H	L	R	s	k	k	v	q	E	L	i	d	y	C	Q	k	c	A	k	n	G	e	-	a	168	

A	168	V	k	L	G	Q	L	L	N	V	C	T	T	N	A	L	A	R	V	M	L	G	K	R	V	F	a	d	G	s	G	197
B	154	V	Q	L	G	Q	L	L	N	V	C	T	T	N	A	L	S	R	V	M	L	G	K	R	V	l	G	d	G	t	G	183
C	178	L	k	L	G	Q	M	M	N	t	C	A	T	N	A	I	A	R	V	M	L	G	R	R	V	v	G	h	a	d	-	206
D	168	V	N	L	G	Q	L	V	N	M	C	v	v	N	A	L	g	R	e	M	I	G	R	R	L	F	G	-	-	-	a	194
E	166	V	N	L	a	Q	L	L	N	L	C	T	v	N	A	L	g	R	V	M	V	G	R	R	V	F	G	d	G	s	G	195
F	166	V	Q	L	G	Q	L	L	N	V	C	A	T	N	A	L	A	R	V	M	L	G	R	R	V	F	G	-	-	-	d	192
G	167	V	p	I	G	Q	M	L	N	V	C	A	T	N	A	L	S	R	V	M	M	G	R	R	V	v	G	h	a	d	G	196
H	165	V	k	L	G	Q	L	L	N	V	C	T	T	N	A	L	A	R	V	M	L	G	R	K	V	F	N	d	G	g	s	194
I	128	V	N	L	G	K	L	L	G	V	C	T	T	N	A	L	A	R	V	M	L	G	R	R	V	F	e	-	G	d	G	156
J	15	I	N	L	G	Q	L	L	G	V	C	T	T	N	A	L	A	R	V	M	L	G	R	R	V	F	G	d	G	s	G	44
K	169	V	d	I	G	R	a	t	f	g	T	T	l	N	l	L	S	n	t	I	f	s	K	d	L	t	N	-	-	-	194	

A	198	d	v	D	P	Q	A	a	E	F	K	S	M	V	V	E	M	M	V	V	A	G	V	F	N	I	G	D	F	I	P	227
B	184	k	s	D	P	K	A	E	E	F	K	d	M	V	L	E	L	M	V	L	T	G	V	F	N	I	G	D	F	V	P	213
C	207	-	-	-	s	K	A	E	E	F	K	A	M	V	V	E	L	M	V	L	A	G	V	F	N	L	G	D	F	I	P	233
D	195	d	a	D	h	K	A	D	E	F	R	S	M	v	t	E	M	M	a	L	A	G	V	F	N	I	G	D	F	V	P	224
E	196	g	d	D	P	K	A	D	E	F	K	S	M	V	V	E	M	M	V	L	A	G	V	F	N	I	G	D	F	I	P	225
F	193	g	i	D	r	s	A	n	E	F	K	d	M	V	V	E	L	M	V	L	A	G	e	F	N	L	G	D	F	I	P	222
G	197	t	n	D	a	K	A	E	E	F	K	A	M	V	V	E	L	M	V	L	S	G	V	F	N	I	G	D	F	I	P	226
H	195	k	s	D	P	K	A	E	E																							

-95- iii

A	228	q	L	n	W	L	D	I	Q	G	V	A	A	K	M	K	K	L	H	A	R	F	D	A	F	L	T	d	I	L	E	257
B	214	a	L	E	c	L	D	L	Q	G	V	A	S	K	M	K	K	L	H	k	R	l	D	n	F	M	S	n	I	L	E	243
C	234	p	L	E	k	L	D	L	Q	G	V	i	A	K	M	K	K	L	H	l	R	F	D	S	F	L	S	k	I	L	g	263
D	225	s	L	D	W	L	D	L	Q	G	V	g	K	M	K	R	L	H	k	R	F	D	A	F	L	S	s	I	L	k	254	
E	226	s	L	E	W	L	D	L	Q	G	V	A	S	K	M	K	K	L	H	k	R	F	D	d	F	L	T	a	I	V	E	255
F	223	v	L	D	l	f	D	L	Q	G	I	T	k	K	M	K	K	L	H	v	R	F	D	S	F	L	S	k	I	V	E	252
G	227	f	L	E	p	L	D	L	Q	G	V	A	S	K	M	K	K	L	H	A	R	F	D	A	F	L	T	e	I	V	r	256
H	225	v	L	g	W	f	D	V	Q	G	I	v	g	K	M	K	K	L	H	A	R	F	D	A	F	L	n	t	I	L	E	254
I	187	v	L	D	W	f	D	L	Q	G	I	A	g	K	M	K	K	L	H	A	R	F	D	k	F	L	n	g	I	L	E	216
J	75	a	L	D	W	f	D	L	Q	G	I	T	A	K	M	K	K	V	H	A	R	F	D	A	F	L	d	a	I	L	E	104
K	225	f	L	E	k	I	D	p	Q	G	I	k	r	R	M	t	n	n	f	T	K	F	l	g	l	I	S	g	L	I	D	254

A	258	E	H	K	g	k	-	-	-	-	i	f	g	e	m	k	D	L	L	S	T	L	I	S	L	K	n	d	d	a	282		
B	244	E	H	K	s	v	a	-	-	-	-	h	q	q	n	g	g	D	L	L	S	i	L	I	S	L	K	-	d	n	c	268	
C	264	D	H	K	i	N	s	s	-	d	e	t	k	g	H	s	-	D	L	L	n	m	L	I	S	L	K	d	a	d	d	291	
D	255	E	H	e	m	N	g	-	-	-	-	q	d	q	K	h	t	D	M	L	S	T	L	I	S	L	K	g	t	d	l	280	
E	256	D	H	K	k	G	s	-	-	-	-	g	t	a	g	h	v	D	M	L	T	T	L	L	S	L	K	-	e	d	a	280	
F	253	E	H	K	t	a	p	-	-	-	-	g	g	l	g	h	t	D	L	L	S	T	L	I	S	L	K	d	d	a	d	278	
G	257	E	R	c	h	G	q	i	-	n	n	s	g	a	H	q	d	D	L	L	S	T	L	I	S	f	K	g	l	d	d	285	
H	255	E	H	K	c	v	n	n	q	h	t	t	l	s	K	d	v	D	f	L	S	T	L	I	r	L	K	d	n	g	a	284	
I	217	D	R	K	s	N	g	s	n	-	-	g	a	e	Q	y	v	D	L	L	S	v	L	I	S	L	Q	d	s	n	i	244	
J	105	E	H	K	s	N	g	s	r	-	-	g	a	k	Q	h	v	D	L	L	S	m	L	I	S	L	Q	d	n	n	i	132	
K	255	D	R	l	k	e	r	n	-	-	-	l	r	d	n	a	n	i	D	V	L	d	A	L	L	n	I	s	q	e	n	p	282

A	283	D	N	d	g	-	-	G	K	L	T	D	T	E	I	K	A	L	L	N	L	F	v	A	G	T	D	T	S	S	310		
B	269	D	G	-	-	e	-	G	G	K	f	S	a	T	E	I	K	A	L	L	L	d	L	F	T	A	G	T	D	T	S	296	
C	292	a	e	-	-	-	-	G	G	R	L	T	D	v	E	I	K	A	L	L	N	L	F	A	A	G	T	D	T	S	318		
D	281	D	G	-	-	d	-	G	G	S	L	T	D	T	E	I	K	A	L	L	N	M	F	T	A	G	T	D	T	S	A	308	
E	281	D	G	-	-	e	-	G	G	K	L	T	D	T	E	I	K	A	L	L	N	M	F	T	A	G	T	D	T	S	S	308	
F	279	i	e	-	-	-	-	G	G	K	L	T	D	T	E	I	K	A	L	L	N	L	F	A	A	G	T	D	T	S	S	305	
G	286	g	d	-	-	-	-	G	S	R	L	T	D	T	E	I	K	A	L	L	N	L	l	-	-	-	D	T	T	S	308		
H	285	D	m	d	c	e	e	G	K	L	T	D	T	E	I	K	A	L	L	N	L	F	T	A	G	T	D	T	S	S	314		
I	245	D	G	g	d	e	-	G	t	K	L	T	D	T	E	I	K	A	L	L	N	L	F	i	A	G	T	D	T	S	S	274	
J	133	D	G	-	e	s	-	G	a	K	L	T	D	T	E	I	K	A	L	L	N	L	F	T	A	G	T	D	T	S	S	161	
K	283	E	e	-	-	-	-	-	-	-	-	I	d	r	N	Q	I	e	q	L	c	L	d	L	F	A	A	G	T	D	T	S	306

A	311	S	T	V	E	W	A	I	A	E	L	I	R	N	P	K	I	L	a	Q	A	Q	Q	E	I	D	k	V	V	G	R	340
B	297	S	T	t	E	W	A	I	A	E	L	I	R	H	P	K	I	L	a	Q	v	Q	Q	E	M	D	s	V	V	G	R	326
C	319	S	T	V	E	W	C	I	A	E	L	V	R	H	P	e	I	L	a	Q	v	Q	k	E	L	D	s	V	V	G	K	348
D	309	S	T	V	D	W	A	I	A	E	L	I	R	H	P	d	I	M	v	K	A	Q	E	E	L	D	i	V	V	G	R	338
E	309	S	T	V	E	W	A	I	A	E	L	I	R	H	P	H	M	L	a	R	v	Q	k	E	L	D	d	f	V	G	H	338
F	306	S	T	V	E	W	A	I	A	E	L	I	R	H	P	Q	I	L	k	Q	A	R	E	E	I	D	a	V	V	G	Q	335
G	309	S	T	V	E	W	A	V	A	E	L	L	R	H	P	K	t	L	a	Q	v	R	Q	E	L	D	s	V	V	G	K	338
H	315	S	T	V	E	W	A	I	A	E	L	L	R	N	P	K	I	L	n	Q	A	Q	Q	E	L	D	l	V	V	G	Q	344
I	275	S	T	V	E	W	A	M	A	E	L	I	R	N	P	K	L	L	v	Q	A	Q	E	E	L	D	r	V	V	G	p	304
J	162	S	T	V	E	W	A	I	A	E	L	I	R	N	P	e	V	L	v	Q	A	Q	Q	E	L	D	r	V	V	G	p	191
K	307	n	T	L	E	W	A	M	A	E	L	L	Q	N	P	H	t	L	g	K	A	Q	E	E	L	a	q	V	I	G	K	336

-96-
iv

A	341	d	R	L	V	g	E	L	D	L	a	Q	L	t	Y	L	E	A	I	V	K	E	T	F	R	L	H	P	S	T	P	370
B	327	d	R	L	I	A	E	A	D	I	p	N	L	t	Y	f	Q	A	V	I	K	E	v	F	R	L	H	P	S	T	P	356
C	349	n	R	V	V	k	B	A	D	L	a	g	L	P	P	L	Q	A	V	V	K	E	N	F	R	L	H	P	S	T	P	378
D	339	d	R	p	V	n	E	S	D	I	a	Q	L	P	Y	L	Q	A	V	I	K	E	N	F	R	L	H	P	p	T	P	368
E	339	d	R	L	V	T	E	S	D	I	p	N	L	P	Y	L	Q	A	V	I	K	E	T	F	R	L	H	P	S	T	P	368
F	336	d	R	L	V	T	E	l	D	L	s	Q	L	t	Y	L	Q	A	L	V	K	E	v	F	R	L	H	P	S	T	P	365
G	339	n	R	L	V	S	E	T	D	L	n	Q	L	P	Y	L	Q	A	V	V	K	E	T	F	R	L	H	P	p	T	P	368
H	345	n	Q	L	V	T	E	S	D	L	t	d	L	P	P	L	Q	A	I	V	K	E	T	F	R	L	H	P	S	T	P	374
I	305	n	R	f	V	T	E	S	D	L	p	Q	L	t	F	L	Q	A	V	I	K	E	T	F	R	L	H	P	S	T	P	334
J	192	s	R	L	V	T	E	S	D	L	p	Q	L	a	P	L	Q	A	V	I	K	E	T	F	R	L	H	P	S	T	P	221
K	337	g	K	q	V	e	E	A	D	V	g	r	L	P	Y	L	r	C	I	V	K	E	T	l	R	I	H	P	A	A	P	366

A	371	L	S	L	P	R	I	A	S	E	S	C	E	I	N	G	Y	f	I	P	K	G	S	T	L	L	L	N	V	W	A	400
B	357	L	S	L	P	R	V	A	n	E	S	C	E	I	N	G	Y	h	I	P	K	N	T	T	L	L	V	N	V	W	A	386
C	379	L	S	L	P	R	I	A	h	E	S	C	E	V	N	G	Y	l	I	P	K	G	S	T	L	L	V	N	V	W	A	408
D	369	L	S	L	P	H	I	A	S	E	S	C	E	I	N	G	Y	h	I	P	K	G	S	T	L	L	t	N	I	W	A	398
E	369	L	S	L	P	R	M	A	A	E	S	C	E	I	N	G	Y	h	I	P	K	G	S	T	L	L	V	N	V	W	A	398
F	366	L	S	L	P	R	I	S	S	E	S	C	E	V	d	G	Y	y	I	P	K	G	S	T	L	L	V	N	V	W	A	395
G	369	L	S	L	P	R	L	A	e	d	d	C	E	I	d	G	Y	l	I	P	K	G	S	T	L	L	V	N	V	W	A	398
H	375	L	S	L	P	R	M	g	A	Q	g	C	E	I	N	G	Y	f	I	P	K	G	A	T	L	L	V	N	V	W	A	404
I	335	L	S	L	P	R	M	A	A	E	d	C	E	I	N	G	Y	y	V	s	e	G	S	T	L	L	V	N	V	W	A	364
J	222	L	S	L	P	R	M	A	S	E	g	C	E	I	N	G	Y	s	I	P	K	G	S	T	L	L	V	N	V	W	S	251
K	367	L	l	I	P	R	k	v	e	E	d	v	E	L	s	t	Y	i	I	P	K	d	S	q	V	L	V	N	V	W	A	396

A	401	I	A	R	D	P	n	a	W	A	D	P	L	E	F	R	P	E	R	F	L	P	G	G	E	K	P	k	V	D	V	430
B	387	I	A	R	D	P	e	V	W	A	D	P	L	E	F	K	P	E	R	F	L	P	G	G	E	K	P	N	V	D	V	416
C	409	I	A	R	D	P	n	V	W	d	E	P	L	E	F	R	P	E	R	F	L	k	G	G	E	K	P	N	V	D	V	438
D	399	I	A	R	D	P	d	q	W	S	D	P	L	a	F	K	P	E	R	F	L	P	G	G	E	K	s	G	V	D	V	428
E	399	I	S	R	D	P	a	e	W	A	D	P	L	E	F	K	P	E	R	F	L	P	G	G	E	K	P	N	V	D	I	428
F	396	I	A	R	D	P	k	M	W	A	D	P	L	E	F	R	P	s	R	F	L	P	G	G	E	K	P	G	a	D	V	425
G	399	I	A	R	D	P	k	V	W	A	D	P	L	E	F	R	P	E	R	F	L	t	G	G	E	K	a	d	V	D	V	428
H	405	I	A	R	D	P	n	V	W	T	n	P	L	E	F	n	P	h	R	F	L	P	G	G	E	K	P	N	V	D	I	434
I	365	I	A	R	D	P	n	a	W	A	n	P	L	D	F	n	P	t	R	F	L	a	G	G	E	K	P	N	V	D	V	394
J	252	I	A	R	D	P	s	I	W	A	D	P	L	E	F	R	P	a	R	F	L	P	G	G	E	K	P	N	V	D	V	281
K	397	I	g	R	n	s	d	L	W	e	n	P	L	v	F	K	P	E	R	F	w	e	s	-	-	-	-	e	I	D	I	422

A	431	R	G	N	D	F	E	V	I	P	F	G	A	G	R	R	I	C	A	G	M	n	L	G	I	R	M	V	Q	L	M	460
B	417	K	G	N	D	F	E	L	I	P	F	G	A	G	R	R	I	C	A	G	L	S	L	G	L	R	M	V	Q	L	M	446
C	439	R	G	N	D	F	E	L	I	P	F	G	A	G	R	R	I	C	A	G	M	S	L	G	I	R	M	V	Q	L	L	468
D	429	K	G	s	D	F	E	L	I	P	F	G	A	G	R	R	I	C	A	G	L	S	L	G	L	R	t	I	Q	f	L	458
E	429	R	G	N	D	F	E	V	I	P	F	G	A	G	R	R	I	C	A	G	M	S	L	G	L	R	M	V	H	L	M	458
F	426	R	G	N	D	F	E	V	I	P	F	G	A	G	R	R	I	C	A	G	M	S	L	G	L	R	M	V	Q	L	L	455
G	429	K	G	N	D	F	E	V	I	P	F	G	A	G	R	R	I	C	A	G	V	g	L	G	I	R	M	V	Q	L	L	458
H	435	K	G	N	D	F	E	V	I	P	F	G	A	G	R	R	I	C	S	G	M	S	L	G	I	R	M	V	H	L	L	464
I	395	K	G	N	D	F	E	V	I	P	F	G	A	G	R	R	I	C	A	G	M	S	L	G	I	R	M	V	Q	L	V	424
J	282	R	G	N	D	F	E	V	I	P	F	G	A	G	R	R	I	C	A	G	M	S	L	G	L	R	M	V	Q	L	s	311
K	423	R	G	r	D	F	E	L	I	P	F	G	A	G	R	R	I	C	p	G	L	p	L	a	M	R	M	I	p	V	a	452

TABLE 8

Percentage of sequence similarity between F3'H sequence of petunia OGR-38 and F3'H sequences from other species and other P450 molecules

	Species / Clone	Number of nucleotides (nt)	Number of amino acids (aa)	%similarity to OGR-38 / no. nt (area of similarity)	%similarity to OGR-38 / no. aa (area of similarity)
5	Petunia OGR-38	1789nt	512aa		
	Snapdragon	1711nt	512aa	69.0% /1573nt	72.2% /507aa
	F3'H cDNA			(19-1578)	(1-504)
	Arabidopsis partial	971nt	270aa	64.7% /745nt	63.7% /248aa
10	F3'H cDNA			(854-1583)	(269-510)
	Arabidopsis Tt7 coding region	1774nt	513aa	65.4% /1066nt	67.1% /511aa
	Carnation	1745nt	496aa	67.3% /1555nt	71.5% /488aa
	F3'H cDNA			(28-1571)	(17-503)
15	Rose	1748nt	513aa	64.7% /1651nt	72.7% /509aa
	F3'H cDNA			(56-1699)	(7-510)
	Gentian	1667nt	476aa	68.3% /1519nt	71.8% /475aa
	partial F3'H cDNA			(170-1673)	(40-510)
	Morning Glory	1824nt	517aa	69.6% /869nt	74.8% /515aa
20	F3'H cDNA			(60-1000)	(3-510)
	Chrysanthemum	1660nt	508aa	68.5% /1532nt	73.6% /511aa
	F3'H cDNA			(50-1580)	(1-510)
	Lisianthus	1214nt	363aa	71.4% /1087nt	74.6% /362aa
	partial F3'H cDNA			(520-1590)	(160-510)
25	Torenia	1815nt	508aa	63.6% /1694nt	67.4% /515aa
	F3'H cDNA			(90-1780)	(1-510)
	Petunia Hf1	1812nt	508aa	58.9% /1471nt	49.9% /513aa
	cDNA			(29-1474)	(1-511)
	Petunia Hf2	1741nt	508aa	58.9% /1481nt	49.1% /511aa
30	cDNA			(37-1498)	(3-510)
	Petunia 651	1716nt	496aa	53.5% /1284nt	38.0% /502aa
	cDNA			(50-1309)	(7-503)
	Mung Bean	1766nt	505aa	56.0% /725nt	29.2% /511aa
35	C4H cDNA			(703-1406)	(1-503)

TABLE 9

Percentage of sequence similarity between F3'H sequence of Snapdragon and F3'H sequences from other species and other P450 molecules

5	Species / Clone	Number of nucleotides (nt)	Number of amino acids (aa)	%similarity to snapdragon/no. nt	%similarity to snapdragon/no. aa
	Snapdragon	1711nt	512aa		
	Petunia OGR-38	1789nt	512aa	69.0% /1573nt	72.2% /507aa
	F3'H cDNA				
	Arabidopsis	971nt	270aa	64.5% /740nt	60.4% /240aa
10	partial F3'H cDNA				
	Carnation	1745nt	496aa	66.7% /1455nt	68.4% /487aa
	F3'H cDNA				
	Torenia	1815nt	508aa	67.6% /1603nt	70.3% /505aa
	F3'H cDNA				
15	Rose	1748nt	513aa	67.2% /1507nt	68.9% /502aa
	F3'H cDNA				
	Petunia Hf1	1812nt	508aa	57.3% /1563nt	49.3% /491aa
	cDNA				
	Petunia Hf2	1741nt	508aa	57.7% /1488nt	47.8% /508aa
20	cDNA				
	Petunia 651	1716nt	496aa	54.4% /1527nt	39.0% /493aa
	cDNA				
	Mung Bean	1766nt	505aa	50.6% /1344nt	32.0% /490aa
25	C4H cDNA				

TABLE 10

Percentage of sequence similarity between F3'H sequence of Arabidopsis and F3'H sequences from other species and other P450 molecules

5	Species / Clone	Number of nucleotides (nt)	Number of amino acids (aa)	%similarity to Arabidopsis/no. nt	%similarity to Arabidopsis/no. aa
	Arabidopsis	971nt	270aa		
	Petunia OGR-38	1789nt	512aa	64.7% /745nt	63.7% /248aa
	F3'H cDNA				
	Snapdragon	1711nt	512aa	64.5%/740nt	60.4%/240aa
10	F3'H cDNA				
	Carnation	1745nt	496aa	64.7% /782nt	60.6%/241aa
	F3'H cDNA				
	Rose	1748nt	513aa	68.5%/739nt	63.7%/248aa
	F3'H cDNA				
15	Petunia 651	1716nt	496aa	57.0%/521nt	40.5%/227aa
	cDNA				
	Petunia Hf1	1812nt	508aa	58.2% /632nt	46.5% /243aa
	cDNA				
	Petunia Hf2	1741nt	508aa	57.4% /632nt	46.1%/243aa
20	cDNA				

TABLE 11

Percentage of sequence similarity between F3'H sequence of Rose and F3'H sequences
from other species and other P450 molecules

5	Species / Clone	Number of nucleotides (nt)	Number of amino acids (aa)	%similarity to Rose	
				/ no. nt	/ no. aa
	Rose	1748bp	513aa		
	Petunia OGR-38	1789bp	512aa	64.7% /1651nt	72.7% /509aa
	F3'H cDNA				
	Snapdragon	1711bp	512aa	67.2% /1507	68.9% /502aa
10	F3'H cDNA				
	Carnation	1745bp	496aa	67.4% /1517nt	72.6% /486aa
	F3'H cDNA				
	Arabidopsis	971bp	270aa	68.5% /739nt	63.7% /248aa
	partial F3'H cDNA				
15	Petunia 651	1716bp	496aa	53.1% /1182nt	37.8% /502aa
	cDNA				
	Petunia Hf1	1812bp	506aa	57% /1366nt	49.9% /503aa
	cDNA				
	Petunia Hf2	1741bp	508aa	57.3% /1331nt	49.1% /505aa
20	cDNA				
	Mung Bean	1766bp	505aa	52.4% /1502nt	32.0% /510aa
	C4H cDNA				

- 102 -

TABLE 12

Percentage of sequence similarity between coding region of Arabidopsis tt7 genomic sequence and F3'H cDNA sequences from other species and other P450 molecules

5	Species / Clone	Number of nucleotides (nt)	Number of amino acids (aa)	%similarity to Arabidopsis tt7 / no. nt	%similarity to Arabidopsis tt7 / no. aa
	Arabidopsis Tt7	1774nt	513aa		
	coding region				
	Petunia OGR-38	1789nt	512aa	65.4 % /1066nt	67.1 %/511aa
	F3'H cDNA				
10	Snapdragon	1711nt	512aa	62.7 %/990nt	64.9 %/504aa
	F3'H cDNA				
	Carnation	1745nt	496aa	63.2 %/1050nt	65.9 %/495aa
	F3'H cDNA				
	Rose	1748nt	513aa	65.5 %/1076nt	68 %/512aa
15	F3'H cDNA				
	Petunia 651	1716nt	496aa	56.5 %/990nt	36.5 %/502aa
	cDNA				
	Petunia Hf1	1812nt	506aa	56.8 %/995nt	47.5 %/509aa
	F3'H cDNA				
20	Petunia Hf2	1741nt	508aa	55.2 %/1063nt	46.8 %/509aa
	F3'H cDNA				

25 Those skilled in the art, will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications. The invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations of any two or more

30 of said steps or features.

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. *J. Mol. Biol.* **215**: 403-410, 1990.

Ashikari, T., Kiuchi-Goto, N., Tanaka, Y., Shibano, Y., Amachi, T., and Yoshizumi, H. *Appl. Microbiol. Biotechnol.* **30**: 515-520, 1989.

Baird, W.V. and Meagher, R. B. *EMBO J.* **6**, 3223-3231, 1987.

Bechtold, N., Ellis, J. and Pelletier, G. C.R. Acad. Sci. Paris, Sciences de la vie **316**: 1194-1199, 1993.

Bethesda Research Laboratories. BRL pUC host: *E. coli* DH5 α [™] competent cells. *Bethesda Res. Lab. Focus.* **8(2)**: 9, 1986.

Brugliera, F., Holton, T.A., Stevenson, T.W., Farcy, E., Lu, C-Y and Cornish, E.C. *Plant J.* **5(1)**: 81-92, 1994.

Church, G.M. and Gilbert, W. *PNAS USA*, **81**: 1991-1995.

Chomczynski, P. and Sacchi, N. *Anal Biochem.* **162**: 156-159.

Comai, L., Moran, P. and Maslyar, D., *Plant Mol. Biol.* **15**: 373-381, 1990.

Cornu, A., Farcy, E., Maizonnier, D., Haring, M., Veerman, W. and Gerats, A.G.M., In: *Genetic maps - Locus maps of complex genomes*. 5th edition, Stephen J. O'Brien (ed.), Cold Spring Harbor Laboratory Press, USA, 1990.

Davies *et al.*, *Plant Science*, **95**: 67-77, 1993.

D'Alessio *et al.*, *Focus*, **14**: 76-79, 1992

De Greve, H., Dhaese, P., Seurinck, J., Lemmers, M., Van Montagu, M and Schell, J. J. *Mol Appl Genet.* **1**: 499-511.

Dellaporta, S.J., Wood, J. and Hick, J.B., *Plant Mol. Biol. Rep.* **1**: 19-21, 1983.

Del Sal, G., Manfioletti, G. and Schneider, C. *Biotechniques*, **7(5)**: 514-519, 1989.

Doodeman, M., Gerats, A.G.M., Schram, A.W., De Vlaming, P. and Bianchi, F., *Theor. Appl. Genet.* **67**: 357-366, 1984.

Dooner, H.K., Robbins, T.R. and Jorgensen, R.A. *Ann. Rev. Genet.* **25**: 173-199, 1991.

Ebel, J. and Hahlbrock, K., In: *The Flavonoids: Advances in Research Since 1980*. Harbourne, J.B. (ed.), Academic Press, New York, USA, 641-679, 1988.

Forkmann, G. and Stotz, G. *Z. Naturforsch.* **36c**:411-416, 1981.

Forkmann, G. *Plant Breeding* **106**: 1-26, 1991.

Franck, A., Guilley, H., Jonard, G. Richards, K. and Hirth, L. *Cell*, **21**, 285-294, 1980.

Frohman, M.A., Dush, M.K., Martin, G.R. *Proc. Natl. Acad. Sci. USA* **85**: 8998-9002, 1988.

Gamborg, O.L., Miller, R.A. and Ojima, K., *Exp. Cell Res.* **50**: 151-158, 1968.

Garfinkel, D.J. and Nester, E.W. *J. Bacteriol.* **144**:732-743, 1980.

Gleave, A.P. *Plant Molecular Biology* 20: 1203-1207, 1992.

Guilley, H., Dudley, R.K., Jonard, G., Balazs, E. and Richards, K.E. *Cell*, 30, 763-773, 1982.

Hahlbrock, K. and Grisebach, H., *Annu. Rev. Plant Physiol.* 30: 105-130, 1979.

Hanahan, D., *J. Mol. Biol.* 166: 557, 1983.

Haughn, G.W. and Somerville, C. *Molecular and General Genetics* 204: 430-434, 1986.

Holton, T.A., Brugliera, F. Lester, D.R., Tanaka, Y., Hyland, C.D., Menting, J.G.T., Lu, C., Farcy, E., Stevenson, T.W. and Cornish, E.C., *Nature* , 366, 276-279, 1993.

Holton, T.A. and Cornish, E.C. *Plant Cell*, 7: 1071-1083, 1995.

Inoue, H., Nojima, H. and Okayama, H. *Gene*, 96: 23-28, 1990.

Ito, H., Fukuda, Y., Murata, K. and Kimura, A. *J. Bacteriol.* 153: 163-168, 1983.

Jefferson, R.A. *Plant Mol. Biol. Rep.* 5: 387-405, 1987.

Jefferson, R.A., Kavanagh, T.A., and Bevan, M.W. *EMBO J.* 6: 3901-3907, 1987

Koornneef, M., Luiten, W., de Vlaming, P. and Schram, A.W. *Arabidopsis Information Service* 19: 113-115, 1982.

Kozak, M. *J. Cell. Biol.* 108: 229, 1989.

Lander, E.S., Green, P., Abrahamson, J., Barlow, A., Day, M.J., Lincoln, S.E. and Newberg, L. *Genomics*, **121**, 185-199, 1987.

Lazo, G.R., Pascal, A.S. and Ludwig, R.A., *Bio/technology*, **9**: 963-967, 1991.

Ledger, S.E., Delores, S.C. and Given, N.K. *Plant Cell Reports*, **10**: 195-199, 1991.

Liang, P. and Pardee, A.B. *Science*, **257**: 967-971, 1992

Liang, P., Averboukh, L. and Pardee, A.B. *Nucl. Acids Res.* **21**: 3269-3275, 1993

Marchuk, D., Drumm, M., Saulino, A., Collins, F.S. *Nucl. Acids Res.* **19**: 1154, 1990

Markham, K. R., *Techniques of flavonoid identification*. London: Academic Press, 1982.

Markham, K. R and Offman, D.J. *Phytochem.*, **34**: 679-685.

Martin, C. and Gerats, T. In: *The molecular biology of flowering*. (Jordan, B.R. ed), UK, CAB International, 219-255, 1993.

McLean, M., Gerats, A.G.M., Baird, W.V. and Meagher, R.B. *J. Heredity* **81**: 341-346, 1990.

Merrifield, J. *Am. Chem. Soc.* **85**: 2149, 1964.

Mizutani, M., Ward, E., DiMaio, J., Ohta, D., Ryals, J. and Sato, R. *Biochem. Biophys. Res. Commun.* **190**: 875-880, 1993.

Murashige, T. and Skoog, F., *Physiol. Plant*, **15**: 73-97, 1962.

Newman, T., de Bruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M. *Plant Physiol.* **106**: 1241-1255, 1994.

Pearson, W.R. and Lipman, D.J., *Proc. Natl. Acad. Sci. USA* **85**: 2444-2448, 1988.

Sambrook, J., Fritsch, E.F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual* (2nd edition). Cold Spring Harbor Laboratory Press, USA, 1989.

Schenk, R.U. and Hilderbrandt, A.C., *Can. J. Bot.* **50**: 199-204, 1972.

Schram, A.W., Jonsson, L.M.V. and Bennink, G.J.H., Biochemistry of flavonoid synthesis in *Petunia hybrida*. In: *Petunia* Sink, K.C. (ed.), Springer-Verlag, Berlin, Germany, pp 68-75, 1984.

Stafford, H.A., *Flavonoid Metabolism*. CRC Press, Inc. Boca Raton, Florida, USA, 1990.

Stotz, G. and Forkmann, G. *Z. Naturforsch* **37c**: 19-23, 1982.

Tabak, A.J.H., Meyer, H. and Bennink, G.J.H., *Planta* **139**, 67-71, 1978.

Tanaka, Y., Ashikari, T., Shibano, Y., Amachi, T., Yoshizumi, H. and Matsubara, H. *J. Biochem.* **103**: 954-961, 1988.

Tanaka, Y., Yonekura, K., Fukuchi-Mizutani, M., Fukui, Y., Fujiwara, H., Ashikari, T. and Kusumi, T. *Plant Cell Physiol.* **37(5)**: 711-716, 1996.

Turpen, T.H. and Griffith, O.M. *BioTechniques*, **4**: 11-15, 1986.

van Tunen A.J. and Mol J.N.M. In: *Plant Biotechnology* (Grierson, D. ed.) Glasgow:

Blackie, 2: 9-31, 1990.

Wiering, H. and de Vlaming, P., Inheritance and Biochemistry of Pigments. In: *Petunia*
Sink, K.C. (ed.), Springer-Verlag, Berlin, Germany, pp 49-65, 1984.

Wallroth, M., Gerats, A.G.M., Rogers, S.G., Fraley, R.T. and Horsch, R.B., *Mol. Gen. Genet.* 202: 6-15, 1986.

- 109 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: (OTHER THAN US): FLORIGENE LIMITED

(US ONLY): Filippa BRUGLIERA, Timothy Albert HOLTON, Michael
Zenon MICHAEL(ii) TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID
PATHWAY ENZYMES AND USES THEREFOR

(iii) NUMBER OF SEQUENCES: 40

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: DAVIES COLLISON CAVE

(B) STREET: 1 LITTLE COLLINS STREET

(C) CITY: MELBOURNE

(D) STATE: VICTORIA

(E) COUNTRY: AUSTRALIA

(F) ZIP: 3000

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:

- 110 -

(B) FILING DATE: 28-FEB-1997

(C) CLASSIFICATION:

(vi) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PN8386

(B) FILING DATE: 28-FEB-1997

(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: HUGHES, DR E JOHN L

(C) REFERENCE/DOCKET NUMBER: EJH/AF

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: +61 3 9254 2777

(B) TELEFAX: +61 3 9254 2770

(C) TELEX: AA 31787

- 111 -

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1789 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 50..1586

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCAGGAATTG GTGAACCCCA TAGAAGTAAA ATACTCCTAT CTTTATTTC ATG GAA	55
Met Glu	
1	
ATC TTA AGC CTA ATT CTG TAC ACC GTC ATT TTC TCA TTT CTT CTA CAA	103
Ile Leu Ser Leu Ile Leu Tyr Thr Val Ile Phe Ser Phe Leu Leu Gln	
5 10 15	
TTC ATT CTT AGA TCA TTT TTC CGT AAA CGT TAC CCT TTA CCA TTA CCA	151
Phe Ile Leu Arg Ser Phe Phe Arg Lys Arg Tyr Pro Leu Pro Leu Pro	
20 25 30	
CCA GGT CCA AAA CCA TGG CCA ATT ATA GGA AAC CTA GTC CAT CTT GGA	199
Pro Gly Pro Lys Pro Trp Pro Ile Ile Gly Asn Leu Val His Leu Gly	
35 40 45 50	
CCC AAA CCA CAT CAA TCA ACT GCA GCC ATG GCT CAA ACT TAT GGA CCA	247
Pro Lys Pro His Gln Ser Thr Ala Ala Met Ala Gln Thr Tyr Gly Pro	
55 60 65	

- 112 -

CTC	ATG	TAT	CTT	AAG	ATG	GGG	TTC	GTA	GAC	GTG	GTG	GTT	GCA	GCC	TCG	295
Leu	Met	Tyr	Leu	Lys	Met	Gly	Phe	Val	Asp	Val	Val	Val	Ala	Ala	Ser	
			70					75					80			
GCA	TCG	GTT	GCA	GCT	CAG	TTC	TTG	AAA	ACT	CAT	GAT	GCT	AAT	TTC	TCG	343
Ala	Ser	Val	Ala	Ala	Gln	Phe	Leu	Lys	Thr	His	Asp	Ala	Asn	Phe	Ser	
			85					90					95			
AGC	CGT	CCA	CCA	AAT	TCT	GGT	GCA	GAA	CAT	ATG	GCT	TAT	AAT	TAT	CAG	391
Ser	Arg	Pro	Pro	Asn	Ser	Gly	Ala	Glu	His	Met	Ala	Tyr	Asn	Tyr	Gln	
			100					105					110			
GAT	CTT	GTT	TTT	GCA	CCT	TAT	GGA	CCT	AGA	TGG	CGT	ATG	CTT	AGG	AAA	439
Asp	Leu	Val	Phe	Ala	Pro	Tyr	Gly	Pro	Arg	Trp	Arg	Met	Leu	Arg	Lys	
			115					120					125		130	
ATT	TGC	TCA	GTT	CAC	CTT	TTC	TCT	ACC	AAG	GCT	TTA	GAT	GAC	TTC	CGC	487
Ile	Cys	Ser	Val	His	Leu	Phe	Ser	Thr	Lys	Ala	Leu	Asp	Asp	Phe	Arg	
				135						140				145		
CAT	GTC	CGC	CAG	GAT	GAA	GTG	AAA	ACA	CTG	ACG	CGC	GCA	CTA	GCA	AGT	535
His	Val	Arg	Gln	Asp	Glu	Val	Lys	Thr	Leu	Thr	Arg	Ala	Leu	Ala	Ser	
				150					155				160			
GCA	GGC	CAA	AAG	CCA	GTC	AAA	TTA	GGT	CAG	TTA	TTG	AAC	GTG	TGC	ACG	583
Ala	Gly	Gln	Lys	Pro	Val	Lys	Leu	Gly	Gln	Leu	Leu	Asn	Val	Cys	Thr	
			165					170					175			
ACG	AAC	GCA	CTC	GCG	CGA	GTA	ATG	CTA	GGT	AAG	CGA	GTA	TTT	GCC	GAC	631
Thr	Asn	Ala	Leu	Ala	Arg	Val	Met	Leu	Gly	Lys	Arg	Val	Phe	Ala	Asp	
			180					185					190			
GGA	AGT	GGC	GAT	GTT	GAT	CCA	CAA	GCG	GCG	GAG	TTC	AAG	TCA	ATG	GTG	679
Gly	Ser	Gly	Asp	Val	Asp	Pro	Gln	Ala	Ala	Glu	Phe	Lys	Ser	Met	Val	
			195					200					205		210	
GTG	GAA	ATG	ATG	GTA	GTC	GCC	GGT	GTT	TTT	AAC	ATT	GGT	GAT	TTT	ATT	727

- 113 -

Val	Glu	Met	Met	Val	Val	Ala	Gly	Val	Phe	Asn	Ile	Gly	Asp	Phe	Ile	
				215				220						225		
CCG	CAA	CTT	AAT	TGG	TTA	GAT	ATT	CAA	GGT	GTA	GCC	GCT	AAA	ATG	AAG	775
Pro	Gln	Leu	Asn	Trp	Leu	Asp	Ile	Gln	Gly	Val	Ala	Ala	Lys	Met	Lys	
			230					235						240		
AAG	CTC	CAC	GCG	CGT	TTC	GAC	GCG	TTC	TTG	ACT	GAT	ATA	CTT	GAA	GAG	823
Lys	Leu	His	Ala	Arg	Phe	Asp	Ala	Phe	Leu	Thr	Asp	Ile	Leu	Glu	Glu	
			245					250						255		
CAT	AAG	GGT	AAA	ATT	TTT	GGA	GAA	ATG	AAA	GAT	TTG	TTG	AGT	ACT	TTG	871
His	Lys	Gly	Lys	Ile	Phe	Gly	Glu	Met	Lys	Asp	Leu	Leu	Ser	Thr	Leu	
			260					265						270		
ATC	TCT	CTT	AAA	AAT	GAT	GAT	GCG	GAT	AAT	GAT	GGA	GGG	AAA	CTC	ACT	919
Ile	Ser	Leu	Lys	Asn	Asp	Asp	Ala	Asp	Asn	Asp	Gly	Gly	Lys	Leu	Thr	
			275					280						285		290
GAT	ACA	GAA	ATT	AAA	GCA	TTA	CTT	TTG	AAC	TTG	TTT	GTA	GCT	GGA	ACA	967
Asp	Thr	Glu	Ile	Lys	Ala	Leu	Leu	Leu	Asn	Leu	Phe	Val	Ala	Gly	Thr	
				295										300		305
GAC	ACA	TCT	TCT	AGT	ACA	GTT	GAA	TGG	GCC	ATT	GCT	GAG	CTT	ATT	CGT	1015
Asp	Thr	Ser	Ser	Ser	Thr	Val	Glu	Trp	Ala	Ile	Ala	Glu	Leu	Ile	Arg	
				310										315		320
AAT	CCA	AAA	ATA	CTA	GCC	CAA	GCC	CAG	CAA	GAG	ATC	GAC	AAA	GTC	GTT	1063
Asn	Pro	Lys	Ile	Leu	Ala	Gln	Ala	Gln	Gln	Glu	Ile	Asp	Lys	Val	Val	
				325										330		335
GGA	AGG	GAC	CGG	CTA	GTT	GGC	GAA	TTG	GAC	CTA	GCC	CAA	TTG	ACA	TAC	1111
Gly	Arg	Asp	Arg	Leu	Val	Gly	Glu	Leu	Asp	Leu	Ala	Gln	Leu	Thr	Tyr	
				340										345		350
TTG	GAA	GCT	ATA	GTC	AAG	GAA	ACC	TTT	CGG	CTT	CAT	CCA	TCA	ACC	CCT	1159
Leu	Glu	Ala	Ile	Val	Lys	Glu	Thr	Phe	Arg	Leu	His	Pro	Ser	Thr	Pro	

- 114 -

355	360	365	370	
CTT TCA CTT CCT AGA ATT GCA TCT GAG AGT TGT GAG ATC AAT GGC TAT				1207
Leu Ser Leu Pro Arg Ile Ala Ser Glu Ser Cys Glu Ile Asn Gly Tyr				
375	380	385		
TTC ATT CCA AAA GGC TCA ACG CTT CTC CTT AAT GTT TGG GCC ATT GCT				1255
Phe Ile Pro Lys Gly Ser Thr Leu Leu Leu Asn Val Trp Ala Ile Ala				
390	395	400		
CGT GAT CCA AAT GCA TGG GCT GAT CCA TTG GAG TTT AGG CCT GAA AGG				1303
Arg Asp Pro Asn Ala Trp Ala Asp Pro Leu Glu Phe Arg Pro Glu Arg				
405	410	415		
TTT TTG CCA GGA GGT GAG AAG CCC AAA GTT GAT GTC CGT GGG AAT GAC				1351
Phe Leu Pro Gly Gly Glu Lys Pro Lys Val Asp Val Arg Gly Asn Asp				
420	425	430		
TTT GAA GTC ATA CCA TTT GGA GCT GGA CGT AGG ATT TGT GCT GGA ATG				1399
Phe Glu Val Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala Gly Met				
435	440	445	450	
AAT TTG GGT ATA CGT ATG GTC CAG TTG ATG ATT GCA ACT TTA ATA CAT				1447
Asn Leu Gly Ile Arg Met Val Gln Leu Met Ile Ala Thr Leu Ile His				
455	460	465		
GCA TTT AAC TGG GAT TTG GTC AGT GGA CAA TTG CCG GAG ATG TTG AAT				1495
Ala Phe Asn Trp Asp Leu Val Ser Gly Gln Leu Pro Glu Met Leu Asn				
470	475	480		
ATG GAA GAA GCA TAT GGG CTG ACC TTA CAA CGG GCT GAT CCA TTG GTT				1543
Met Glu Glu Ala Tyr Gly Leu Thr Leu Gln Arg Ala Asp Pro Leu Val				
485	490	495		
GTG CAC CCA AGG CCT CGC TTA GAA GCC CAA GCG TAC ATT GGG T				1586
Val His Pro Arg Pro Arg Leu Glu Ala Gln Ala Tyr Ile Gly				
500	505	510		

- 115 -

GAGCAGCAAC AGCCCATGGA GATAACATGA GTGTAAATG TATGAGTCTC CATATCTTGT 1646

TTAGTTTGT TATGCTTTGG ATTTAGTAGT TTTTATATTG ATAGATCAAT GTTTCATTG 1706

TCAGTAAGAA TATCCGTTGC TTGTTTCATT AACTCCAGGT GGACAATAAA AGAAGTAATT 1766

TGTATGAAAA AAAAAAAAAA AAA 1789

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Glu Ile Leu Ser Leu Ile Leu Tyr Thr Val Ile Phe Ser Phe Leu
 1 5 10 15

Leu Gln Phe Ile Leu Arg Ser Phe Phe Arg Lys Arg Tyr Pro Leu Pro
 20 25 30

Leu Pro Pro Gly Pro Lys Pro Trp Pro Ile Ile Gly Asn Leu Val His
 35 40 45

Leu Gly Pro Lys Pro His Gln Ser Thr Ala Ala Met Ala Gln Thr Tyr
 50 55 60

Gly Pro Leu Met Tyr Leu Lys Met Gly Phe Val Asp Val Val Val Ala
 65 70 75 80

Ala Ser Ala Ser Val Ala Ala Gln Phe Leu Lys Thr His Asp Ala Asn
 85 90 95

- 116 -

Phe Ser Ser Arg Pro Pro Asn Ser Gly Ala Glu His Met Ala Tyr Asn

100

105

110

Tyr Gln Asp Leu Val Phe Ala Pro Tyr Gly Pro Arg Trp Arg Met Leu

115

120

125

Arg Lys Ile Cys Ser Val His Leu Phe Ser Thr Lys Ala Leu Asp Asp

130

135

140

Phe Arg His Val Arg Gln Asp Glu Val Lys Thr Leu Thr Arg Ala Leu

145

150

155

160

Ala Ser Ala Gly Gln Lys Pro Val Lys Leu Gly Gln Leu Leu Asn Val

165

170

175

Cys Thr Thr Asn Ala Leu Ala Arg Val Met Leu Gly Lys Arg Val Phe

180

185

190

Ala Asp Gly Ser Gly Asp Val Asp Pro Gln Ala Ala Glu Phe Lys Ser

195

200

205

Met Val Val Glu Met Met Val Val Ala Gly Val Phe Asn Ile Gly Asp

210

215

220

Phe Ile Pro Gln Leu Asn Trp Leu Asp Ile Gln Gly Val Ala Ala Lys

225

230

235

240

Met Lys Lys Leu His Ala Arg Phe Asp Ala Phe Leu Thr Asp Ile Leu

245

250

255

Glu Glu His Lys Gly Lys Ile Phe Gly Glu Met Lys Asp Leu Leu Ser

260

265

270

Thr Leu Ile Ser Leu Lys Asn Asp Asp Ala Asp Asn Asp Gly Gly Lys

275

280

285

Leu Thr Asp Thr Glu Ile Lys Ala Leu Leu Leu Asn Leu Phe Val Ala

- 117 -

290	295	300
Gly Thr Asp Thr Ser Ser Ser Thr Val Glu Trp Ala Ile Ala Glu Leu		
305	310	315 320
Ile Arg Asn Pro Lys Ile Leu Ala Gln Ala Gln Gln Glu Ile Asp Lys		
	325	330 335
Val Val Gly Arg Asp Arg Leu Val Gly Glu Leu Asp Leu Ala Gln Leu		
	340	345 350
Thr Tyr Leu Glu Ala Ile Val Lys Glu Thr Phe Arg Leu His Pro Ser		
	355	360 365
Thr Pro Leu Ser Leu Pro Arg Ile Ala Ser Glu Ser Cys Glu Ile Asn		
	370 375	380
Gly Tyr Phe Ile Pro Lys Gly Ser Thr Leu Leu Leu Asn Val Trp Ala		
385	390	395 400
Ile Ala Arg Asp Pro Asn Ala Trp Ala Asp Pro Leu Glu Phe Arg Pro		
	405	410 415
Glu Arg Phe Leu Pro Gly Gly Glu Lys Pro Lys Val Asp Val Arg Gly		
	420	425 430
Asn Asp Phe Glu Val Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala		
	435	440 445
Gly Met Asn Leu Gly Ile Arg Met Val Gln Leu Met Ile Ala Thr Leu		
	450	455 460
Ile His Ala Phe Asn Trp Asp Leu Val Ser Gly Gln Leu Pro Glu Met		
465	470	475 480
Leu Asn Met Glu Glu Ala Tyr Gly Leu Thr Leu Gln Arg Ala Asp Pro		
	485	490 495

- 118 -

Leu Val Val His Pro Arg Pro Arg L u Glu Ala Gln Ala Tyr Ile Gly

500

505

510

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1745 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 172..1660

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

AAGTTCGGCA CGAGCGTCAC ATTCACACCG TCACATTACT ATTCAAACCA CTCATTTTCT      60

ACCTCTCTTT TCTACCCACC AAAACAAAAC AAAACAAAAA AAAACACATA AAAAAACTCA      120

AAAAAAAATT ATAATGTCAC CCTTAGAGGT AACTTTCTAC ACCATAGTCC T ATG CAC      177
                               Met His
                               1

AAT CTC TAC TAC CTC ATC ACC ACC GTC TTC CGC GGC CAC CAA AAA CCG      225
Asn Leu Tyr Tyr Leu Ile Thr Thr Val Phe Arg Gly His Gln Lys Pro
      5              10              15

CTT CCT CCA GGG CCA CGA CCA TGG CCC ATC GTG GGA AAC CTC CCA CAT      273
Leu Pro Pro Gly Pro Arg Pro Trp Pro Ile Val Gly Asn Leu Pro His
      20              25              30

```

- 119 -

ATG GGC CAG GCA CCG CAC CAG GGC TTA GCA GCC CTG GCG CAA AAG TAT	321
Met Gly Gln Ala Pro His Gln Gly Leu Ala Ala Leu Ala Gln Lys Tyr	
35 40 45 50	
GGC CCT CTA TTG TAT ATG AGA CTG GGG TAC GTG GAC GTT GTT GTG GCC	369
Gly Pro Leu Leu Tyr Met Arg Leu Gly Tyr Val Asp Val Val Val Ala	
55 60 65	
GCC TCA GCG TCT GTA GCG ACC CAG TTT CTT AAG ACA CAT GAC CTA AAT	417
Ala Ser Ala Ser Val Ala Thr Gln Phe Leu Lys Thr His Asp Leu Asn	
70 75 80	
TTT TCG AGT AGG CCA CCG AAT TCG GGG GCT AAA CAC ATT GCT TAT AAC	465
Phe Ser Ser Arg Pro Pro Asn Ser Gly Ala Lys His Ile Ala Tyr Asn	
85 90 95	
TAT CAA GAC CTT GTT TTT GCA CCT TAT GGA CCT AAA TGG CGC ATG CTT	513
Tyr Gln Asp Leu Val Phe Ala Pro Tyr Gly Pro Lys Trp Arg Met Leu	
100 105 110	
AGG AAA ATT TGT TCC TTA CAC ATG TTT TCT TCT AAG GCT TTG GAC GAT	561
Arg Lys Ile Cys Ser Leu His Met Phe Ser Ser Lys Ala Leu Asp Asp	
115 120 125 130	
TTT AGA CTT GTC CGT CAG GAA GAA GTA TCT ATA CTG GTA AAT GCG ATA	609
Phe Arg Leu Val Arg Gln Glu Glu Val Ser Ile Leu Val Asn Ala Ile	
135 140 145	
GCA AAA GCA GGA ACA AAG CCA GTA CAA CTA GGA CAA CTA CTC AAC GTG	657
Ala Lys Ala Gly Thr Lys Pro Val Gln Leu Gly Gln Leu Leu Asn Val	
150 155 160	
TGC ACC ACA AAT GCC TTA TCG AGG GTG ATG CTA GGG AAG CGA GTT CTC	705
Cys Thr Thr Asn Ala Leu Ser Arg Val Met Leu Gly Lys Arg Val Leu	
165 170 175	
GGT GAT GGC ACA GGG AAA AGC GAC CCA AAA GCC GAG GAA TTT AAG GAC	753

- 120 -

Gly	Asp	Gly	Thr	Gly	Lys	Ser	Asp	Pro	Lys	Ala	Glu	Glu	Phe	Lys	Asp	
180						185					190					
ATG	GTG	CTG	GAG	TTA	ATG	GTT	CTC	ACC	GGA	GTT	TTT	AAC	ATT	GGC	GAT	801
Met	Val	Leu	Glu	Leu	Met	Val	Leu	Thr	Gly	Val	Phe	Asn	Ile	Gly	Asp	
195					200				205					210		
TTT	GTA	CCG	GCA	TTG	GAA	TGT	CTA	GAC	TTA	CAA	GGT	GTT	GCA	TCT	AAA	849
Phe	Val	Pro	Ala	Leu	Glu	Cys	Leu	Asp	Leu	Gln	Gly	Val	Ala	Ser	Lys	
				215				220						225		
ATG	AAG	AAA	TTA	CAT	AAA	AGA	CTT	GAT	AAT	TTT	ATG	AGT	AAC	ATT	TTG	897
Met	Lys	Lys	Leu	His	Lys	Arg	Leu	Asp	Asn	Phe	Met	Ser	Asn	Ile	Leu	
			230					235						240		
GAG	GAA	CAC	AAG	AGT	GTT	GCA	CAT	CAA	CAA	AAT	GGT	GGA	GAT	TTG	CTA	945
Glu	Glu	His	Lys	Ser	Val	Ala	His	Gln	Gln	Asn	Gly	Gly	Asp	Leu	Leu	
		245						250						255		
AGC	ATT	TTG	ATA	TCT	TTG	AAG	GAT	AAT	TGT	GAT	GGT	GAA	GGT	GGC	AAG	993
Ser	Ile	Leu	Ile	Ser	Leu	Lys	Asp	Asn	Cys	Asp	Gly	Glu	Gly	Gly	Lys	
		260						265						270		
TTT	AGT	GCC	ACA	GAA	ATT	AAG	GCC	TTG	CTA	TTG	GAT	TTA	TTT	ACA	GCT	1041
Phe	Ser	Ala	Thr	Glu	Ile	Lys	Ala	Leu	Leu	Leu	Asp	Leu	Phe	Thr	Ala	
275					280					285					290	
GGA	ACA	GAC	ACA	TCA	TCT	AGT	ACA	ACT	GAA	TGG	GCC	ATA	GCC	GAA	CTA	1089
Gly	Thr	Asp	Thr	Ser	Ser	Ser	Thr	Thr	Glu	Trp	Ala	Ile	Ala	Glu	Leu	
				295						300				305		
ATT	CGC	CAT	CCA	AAA	ATC	TTA	GCC	CAA	GTT	CAA	CAA	GAA	ATG	GAC	TCA	1137
Ile	Arg	His	Pro	Lys	Ile	Leu	Ala	Gln	Val	Gln	Gln	Glu	Met	Asp	Ser	
			310					315						320		
GTC	GTG	GGC	CGA	GAC	CGA	CTC	ATA	GCC	GAA	GCT	GAC	ATA	CCG	AAC	CTA	1185
Val	Val	Gly	Arg	Asp	Arg	Leu	Ile	Ala	Glu	Ala	Asp	Ile	Pro	Asn	Leu	

- 121 -

325	330	335	
ACC TAC TTC CAA GCC GTA ATC AAA GAG GTT TTC CGA CTT CAC CCG TCC			1233
Thr Tyr Phe Gln Ala Val Ile Lys Glu Val Phe Arg Leu His Pro Ser			
340	345	350	
ACC CCG CTT TCC CTA CCA CGG GTC GCA AAC GAA TCG TGC GAA ATA AAC			1281
Thr Pro Leu Ser Leu Pro Arg Val Ala Asn Glu Ser Cys Glu Ile Asn			
355	360	365	370
GGG TAC CAC ATT CCC AAA AAC ACC ACT TTA TTG GTA AAT GTG TGG GCC			1329
Gly Tyr His Ile Pro Lys Asn Thr Thr Leu Leu Val Asn Val Trp Ala			
375	380	385	
ATC GCA CGC GAC CCT GAG GTT TGG GCC GAC CCG TTA GAG TTT AAA CCC			1377
Ile Ala Arg Asp Pro Glu Val Trp Ala Asp Pro Leu Glu Phe Lys Pro			
390	395	400	
GAA AGA TTT TTG CCG GGC GGC GAA AAG CCC AAT GTG GAT GTG AAA GGA			1425
Glu Arg Phe Leu Pro Gly Gly Glu Lys Pro Asn Val Asp Val Lys Gly			
405	410	415	
AAC GAT TTT GAG CTG ATT CCG TTC GGG GCG GGC CGA CGG ATT TGT GCT			1473
Asn Asp Phe Glu Leu Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala			
420	425	430	
GGG CTG AGT TTG GGC CTG CGT ATG GTC CAG TTG ATG ACA GCC ACT TTG			1521
Gly Leu Ser Leu Gly Leu Arg Met Val Gln Leu Met Thr Ala Thr Leu			
435	440	445	450
GCC CAT ACT TAT GAT TGG GCC TTA GCT GAT GGG CTT ATG CCC GAA AAG			1569
Ala His Thr Tyr Asp Trp Ala Leu Ala Asp Gly Leu Met Pro Glu Lys			
455	460	465	
CTT AAC ATG GAT GAG GCT TAT GGG CTT ACC TTA CAG CGT AAG GTG CCA			1617
Leu Asn Met Asp Glu Ala Tyr Gly Leu Thr Leu Gln Arg Lys Val Pro			
470	475	480	

- 122 -

CTT AAT GGT CCA CCC GAC CCC GTC GGC TTC TCG GCC CGT GTT T 1660

Leu Asn Gly Pro Pro Asp Pro Val Gly Phe Ser Ala Arg Val

485

490

495

AATAATTCCG GGGTTTTTAA AAGCGGGTTA CTTTGTGTTA TGTATTATTC CGTACTAGTT 1720

TGAAAATAAT GGTATTAGAG AAATG 1745

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 496 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met His Asn Leu Tyr Tyr Leu Ile Thr Thr Val Phe Arg Gly His Gln

1

5

10

15

Lys Pro Leu Pro Pro Gly Pro Arg Pro Trp Pro Ile Val Gly Asn Leu

20

25

30

Pro His Met Gly Gln Ala Pro His Gln Gly Leu Ala Ala Leu Ala Gln

35

40

45

Lys Tyr Gly Pro Leu Leu Tyr Met Arg Leu Gly Tyr Val Asp Val Val

50

55

60

Val Ala Ala Ser Ala Ser Val Ala Thr Gln Phe Leu Lys Thr His Asp

65

70

75

80

Leu Asn Phe Ser Ser Arg Pro Pro Asn Ser Gly Ala Lys His Ile Ala

85

90

95

- 123 -

Tyr Asn Tyr Gln Asp Leu Val Phe Ala Pro Tyr Gly Pro Lys Trp Arg

100

105

110

Met Leu Arg Lys Ile Cys Ser Leu His Met Phe Ser Ser Lys Ala Leu

115

120

125

Asp Asp Phe Arg Leu Val Arg Gln Glu Glu Val Ser Ile Leu Val Asn

130

135

140

Ala Ile Ala Lys Ala Gly Thr Lys Pro Val Gln Leu Gly Gln Leu Leu

145

150

155

160

Asn Val Cys Thr Thr Asn Ala Leu Ser Arg Val Met Leu Gly Lys Arg

165

170

175

Val Leu Gly Asp Gly Thr Gly Lys Ser Asp Pro Lys Ala Glu Glu Phe

180

185

190

Lys Asp Met Val Leu Glu Leu Met Val Leu Thr Gly Val Phe Asn Ile

195

200

205

Gly Asp Phe Val Pro Ala Leu Glu Cys Leu Asp Leu Gln Gly Val Ala

210

215

220

Ser Lys Met Lys Lys Leu His Lys Arg Leu Asp Asn Phe Met Ser Asn

225

230

235

240

Ile Leu Glu Glu His Lys Ser Val Ala His Gln Gln Asn Gly Gly Asp

245

250

255

Leu Leu Ser Ile Leu Ile Ser Leu Lys Asp Asn Cys Asp Gly Glu Gly

260

265

270

Gly Lys Phe Ser Ala Thr Glu Ile Lys Ala Leu Leu Leu Asp Leu Phe

275

280

285

Thr Ala Gly Thr Asp Thr Ser Ser Ser Thr Thr Glu Trp Ala Ile Ala

- 124 -

290	295	300
Glu Leu Ile Arg His Pro Lys Ile Leu Ala Gln Val Gln Gln Glu Met		
305	310	315 320
Asp Ser Val Val Gly Arg Asp Arg Leu Ile Ala Glu Ala Asp Ile Pro		
325	330	335
Asn Leu Thr Tyr Phe Gln Ala Val Ile Lys Glu Val Phe Arg Leu His		
340	345	350
Pro Ser Thr Pro Leu Ser Leu Pro Arg Val Ala Asn Glu Ser Cys Glu		
355	360	365
Ile Asn Gly Tyr His Ile Pro Lys Asn Thr Thr Leu Leu Val Asn Val		
370	375	380
Trp Ala Ile Ala Arg Asp Pro Glu Val Trp Ala Asp Pro Leu Glu Phe		
385	390	395 400
Lys Pro Glu Arg Phe Leu Pro Gly Gly Glu Lys Pro Asn Val Asp Val		
405	410	415
Lys Gly Asn Asp Phe Glu Leu Ile Pro Phe Gly Ala Gly Arg Arg Ile		
420	425	430
Cys Ala Gly Leu Ser Leu Gly Leu Arg Met Val Gln Leu Met Thr Ala		
435	440	445
Thr Leu Ala His Thr Tyr Asp Trp Ala Leu Ala Asp Gly Leu Met Pro		
450	455	460
Glu Lys Leu Asn Met Asp Glu Ala Tyr Gly Leu Thr Leu Gln Arg Lys		
465	470	475 480
Val Pro Leu Asn Gly Pro Pro Asp Pro Val Gly Phe Ser Ala Arg Val		
485	490	495

- 125 -

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1711 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 91..1629

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGAATTCCCC CCCCCCACA CCATTCAATG CCTAAGTCCT CCATTGCCG GCCTAATAAC	60
TAAAAGCCCA CTCTTCCGA CCATCTATAC ATG CAA CAC CAA TAT TAT TCT TTA	114
Met Gln His Gln Tyr Tyr Ser Leu	
1 5	
ATT ACG ATG GAT GAT ATT AGC ATA ACC AGC TTA TTG GTG CCA TGT ACT	162
Ile Thr Met Asp Asp Ile Ser Ile Thr Ser Leu Leu Val Pro Cys Thr	
10 15 20	
TTT ATA TTA GGG TTC TTG CTT CTA TAT TCC TTC CTC AAC AAA AAA GTA	210
Phe Ile Leu Gly Phe Leu Leu Leu Tyr Ser Phe Leu Asn Lys Lys Val	
25 30 35 40	
AAG CCA CTG CCA CCT GGA CCG AAG CCA TGG CCC ATC GTC GGA AAT CTG	258
Lys Pro Leu Pro Pro Gly Pro Lys Pro Trp Pro Ile Val Gly Asn Leu	
45 50 55	
CCA CAT CTT GGG CCG AAG CCC CAC CAG TCG ATG GCG GCG CTG GCA CGG	306

- 126 -

Pro	His	Leu	Gly	Pro	Lys	Pro	His	Gln	Ser	Met	Ala	Ala	Leu	Ala	Arg	
				60				65					70			
GTG	CAC	GGC	CCA	TTA	ATT	CAT	CTG	AAG	ATG	GGC	TTT	GTG	CAT	GTG	GTT	354
Val	His	Gly	Pro	Leu	Ile	His	Leu	Lys	Met	Gly	Phe	Val	His	Val	Val	
		75						80					85			
GTG	GCC	TCC	TCA	GCA	TCC	GTT	GCG	GAG	AAA	TTT	CTG	AAG	GTG	CAT	GAC	402
Val	Ala	Ser	Ser	Ala	Ser	Val	Ala	Glu	Lys	Phe	Leu	Lys	Val	His	Asp	
		90					95					100				
GCA	AAC	TTC	TCG	AGC	AGG	CCT	CCC	AAT	TCG	GGT	GCA	AAA	CAC	GTG	GCC	450
Ala	Asn	Phe	Ser	Ser	Arg	Pro	Pro	Asn	Ser	Gly	Ala	Lys	His	Val	Ala	
105						110				115				120		
TAC	AAC	TAT	CAG	GAC	TTG	GTC	TTT	GCT	CCT	TAT	GGC	CCA	CGC	TGG	CGG	498
Tyr	Asn	Tyr	Gln	Asp	Leu	Val	Phe	Ala	Pro	Tyr	Gly	Pro	Arg	Trp	Arg	
			125					130					135			
ATG	CTC	AGG	AAA	ATC	TGT	GCA	CTC	CAC	CTC	TTC	TCC	GCC	AAA	GCC	TTG	546
Met	Leu	Arg	Lys	Ile	Cys	Ala	Leu	His	Leu	Phe	Ser	Ala	Lys	Ala	Leu	
			140					145					150			
AAC	GAC	TTC	ACA	CAC	GTC	AGA	CAG	GAT	GAG	GTG	GGG	ATC	CTC	ACT	CGC	594
Asn	Asp	Phe	Thr	His	Val	Arg	Gln	Asp	Glu	Val	Gly	Ile	Leu	Thr	Arg	
		155					160					165				
GTT	CTA	GCA	GAT	GCA	GGA	GAA	ACG	CCG	TTG	AAA	TTA	GGG	CAG	ATG	ATG	642
Val	Leu	Ala	Asp	Ala	Gly	Glu	Thr	Pro	Leu	Lys	Leu	Gly	Gln	Met	Met	
		170					175					180				
AAC	ACA	TGC	GCC	ACC	AAT	GCA	ATA	GCG	CGT	GTT	ATG	TTG	GGT	CGA	CGC	690
Asn	Thr	Cys	Ala	Thr	Asn	Ala	Ile	Ala	Arg	Val	Met	Leu	Gly	Arg	Arg	
185					190				195				200			
GTG	GTT	GGA	CAC	GCA	GAC	TCA	AAG	GCG	GAG	GAG	TTT	AAG	GCA	ATG	GTA	738
Val	Val	Gly	His	Ala	Asp	Ser	Lys	Ala	Glu	Glu	Phe	Lys	Ala	Met	Val	

- 127 -

	205	210	215	
GTG GAG TTG ATG GTA TTA GCT GGT GTG TTC AAC TTA GGT GAT TTT ATC				786
Val Glu Leu Met Val Leu Ala Gly Val Phe Asn Leu Gly Asp Phe Ile				
	220	225	230	
CCA CCT CTT GAA AAA TTG GAT CTT CAA GGT GTC ATT GCT AAG ATG AAG				834
Pro Pro Leu Glu Lys Leu Asp Leu Gln Gly Val Ile Ala Lys Met Lys				
	235	240	245	
AAG CTT CAC TTG CGT TTC GAC TCG TTC TTG AGT AAG ATC CTT GGA GAC				882
Lys Leu His Leu Arg Phe Asp Ser Phe Leu Ser Lys Ile Leu Gly Asp				
	250	255	260	
CAC AAG ATC AAC AGC TCA GAT GAA ACC AAA GGC CAT TCG GAT TTG TTG				930
His Lys Ile Asn Ser Ser Asp Glu Thr Lys Gly His Ser Asp Leu Leu				
	265	270	275	280
AAC ATG TTA ATT TCT TTG AAG GAC GCT GAT GAT GCC GAA GGA GGG AGG				978
Asn Met Leu Ile Ser Leu Lys Asp Ala Asp Asp Ala Glu Gly Gly Arg				
	285	290	295	
CTC ACC GAC GTA GAA ATT AAA GCG TTG CTC TTG AAC TTG TTT GCT GCA				1026
Leu Thr Asp Val Glu Ile Lys Ala Leu Leu Leu Asn Leu Phe Ala Ala				
	300	305	310	
GGA ACT GAC ACA ACA TCA AGC ACT GTG GAA TGG TGC ATA GCT GAG TTA				1074
Gly Thr Asp Thr Thr Ser Ser Thr Val Glu Trp Cys Ile Ala Glu Leu				
	315	320	325	
GTA CGA CAT CCT GAA ATC CTT GCC CAA GTC CAA AAA GAA CTC GAC TCT				1122
Val Arg His Pro Glu Ile Leu Ala Gln Val Gln Lys Glu Leu Asp Ser				
	330	335	340	
GTT GTT GGT AAG AAT CGG GTG GTG AAG GAG GCT GAT CTG GCC GGA TTA				1170
Val Val Gly Lys Asn Arg Val Val Lys Glu Ala Asp Leu Ala Gly Leu				
	345	350	355	360

- 128 -

CCA	TTC	CTC	CAA	GCG	GTC	GTC	AAG	GAA	AAT	TTC	CGA	CTC	CAT	CCC	TCC	1218
Pro	Phe	Leu	Gln	Ala	Val	Val	Lys	Glu	Asn	Phe	Arg	Leu	His	Pro	Ser	
				365					370					375		
ACC	CCG	CTC	TCC	CTA	CCG	AGG	ATC	GCA	CAT	GAG	AGT	TGT	GAA	GTG	AAT	1266
Thr	Pro	Leu	Ser	Leu	Pro	Arg	Ile	Ala	His	Glu	Ser	Cys	Glu	Val	Asn	
				380				385					390			
GGA	TAC	TTG	ATT	CCA	AAG	GGT	TCG	ACA	CTT	CTT	GTC	AAT	GTT	TGG	GCA	1314
Gly	Tyr	Leu	Ile	Pro	Lys	Gly	Ser	Thr	Leu	Leu	Val	Asn	Val	Trp	Ala	
				395				400					405			
ATT	GCT	CGC	GAT	CCA	AAT	GTG	TGG	GAT	GAA	CCA	CTA	GAG	TTC	CGG	CCT	1362
Ile	Ala	Arg	Asp	Pro	Asn	Val	Trp	Asp	Glu	Pro	Leu	Glu	Phe	Arg	Pro	
				410				415					420			
GAA	CGA	TTC	TTG	AAG	GGC	GGG	GAA	AAG	CCT	AAT	GTC	GAT	GTT	AGA	GGG	1410
Glu	Arg	Phe	Leu	Lys	Gly	Gly	Glu	Lys	Pro	Asn	Val	Asp	Val	Arg	Gly	
				425				430					435		440	
AAT	GAT	TTC	GAA	TTG	ATA	CCG	TTC	GGA	GCG	GGC	CGA	AGA	ATT	TGT	GCA	1458
Asn	Asp	Phe	Glu	Leu	Ile	Pro	Phe	Gly	Ala	Gly	Arg	Arg	Ile	Cys	Ala	
				445				450					455			
GGA	ATG	AGC	TTA	GGA	ATA	CGT	ATG	GTC	CAG	TTG	TTG	ACA	GCA	ACT	TTG	1506
Gly	Met	Ser	Leu	Gly	Ile	Arg	Met	Val	Gln	Leu	Leu	Thr	Ala	Thr	Leu	
				460				465					470			
AAC	CAT	GCG	TTT	GAC	TTT	GAT	TTG	GCG	GAT	GGA	CAG	TTG	CCT	GAA	AGC	1554
Asn	His	Ala	Phe	Asp	Phe	Asp	Leu	Ala	Asp	Gly	Gln	Leu	Pro	Glu	Ser	
				475				480					485			
TTA	AAC	ATG	GAG	GAA	GCT	TAT	GGG	CTG	ACC	TTG	CAA	CGA	GCT	GAC	CCT	1602
Leu	Asn	Met	Glu	Glu	Ala	Tyr	Gly	Leu	Thr	Leu	Gln	Arg	Ala	Asp	Pro	
				490				495					500			
TTG	GTA	GTG	CAC	CCG	AAG	CCT	AGG	TAGGCACCTC	ATGTTTATCA	AACTTAGGAC						1656

- 129 -

Leu Val Val His Pro Lys Pro Arg

505

510

TCATGTTTAG AGAACCTCTT GTTGTTTTAT CAGATTGAAG TGTGATGTCC AAGAC

1711

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 512 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Gln His Gln Tyr Tyr Ser Leu Ile Thr Met Asp Asp Ile Ser Ile
1 5 10 15

Thr Ser Leu Leu Val Pro Cys Thr Phe Ile Leu Gly Phe Leu Leu Leu
20 25 30

Tyr Ser Phe Leu Asn Lys Lys Val Lys Pro Leu Pro Pro Gly Pro Lys
35 40 45

Pro Trp Pro Ile Val Gly Asn Leu Pro His Leu Gly Pro Lys Pro His
50 55 60

Gln Ser Met Ala Ala Leu Ala Arg Val His Gly Pro Leu Ile His Leu
65 70 75 80

Lys Met Gly Phe Val His Val Val Val Ala Ser Ser Ala Ser Val Ala
85 90 95

Glu Lys Phe Leu Lys Val His Asp Ala Asn Phe Ser Ser Arg Pro Pro
100 105 110

- 130 -

Asn Ser Gly Ala Lys His Val Ala Tyr Asn Tyr Gln Asp Leu Val Phe

115

120

125

Ala Pro Tyr Gly Pro Arg Trp Arg Met Leu Arg Lys Ile Cys Ala Leu

130

135

140

His Leu Phe Ser Ala Lys Ala Leu Asn Asp Phe Thr His Val Arg Gln

145

150

155

160

Asp Glu Val Gly Ile Leu Thr Arg Val Leu Ala Asp Ala Gly Glu Thr

165

170

175

Pro Leu Lys Leu Gly Gln Met Met Asn Thr Cys Ala Thr Asn Ala Ile

180

185

190

Ala Arg Val Met Leu Gly Arg Arg Val Val Gly His Ala Asp Ser Lys

195

200

205

Ala Glu Glu Phe Lys Ala Met Val Val Glu Leu Met Val Leu Ala Gly

210

215

220

Val Phe Asn Leu Gly Asp Phe Ile Pro Pro Leu Glu Lys Leu Asp Leu

225

230

235

240

Gln Gly Val Ile Ala Lys Met Lys Lys Leu His Leu Arg Phe Asp Ser

245

250

255

Phe Leu Ser Lys Ile Leu Gly Asp His Lys Ile Asn Ser Ser Asp Glu

260

265

270

Thr Lys Gly His Ser Asp Leu Leu Asn Met Leu Ile Ser Leu Lys Asp

275

280

285

Ala Asp Asp Ala Glu Gly Gly Arg Leu Thr Asp Val Glu Ile Lys Ala

290

295

300

Leu Leu Leu Asn Leu Phe Ala Ala Gly Thr Asp Thr Thr Ser Ser Thr

- 131 -

305	310	315	320
Val Glu Trp Cys Ile Ala Glu Leu Val Arg His Pro Glu Ile Leu Ala	325	330	335
Gln Val Gln Lys Glu Leu Asp Ser Val Val Gly Lys Asn Arg Val Val	340	345	350
Lys Glu Ala Asp Leu Ala Gly Leu Pro Phe Leu Gln Ala Val Val Lys	355	360	365
Glu Asn Phe Arg Leu His Pro Ser Thr Pro Leu Ser Leu Pro Arg Ile	370	375	380
Ala His Glu Ser Cys Glu Val Asn Gly Tyr Leu Ile Pro Lys Gly Ser	385	390	395
Thr Leu Leu Val Asn Val Trp Ala Ile Ala Arg Asp Pro Asn Val Trp	405	410	415
Asp Glu Pro Leu Glu Phe Arg Pro Glu Arg Phe Leu Lys Gly Gly Glu	420	425	430
Lys Pro Asn Val Asp Val Arg Gly Asn Asp Phe Glu Leu Ile Pro Phe	435	440	445
Gly Ala Gly Arg Arg Ile Cys Ala Gly Met Ser Leu Gly Ile Arg Met	450	455	460
Val Gln Leu Leu Thr Ala Thr Leu Asn His Ala Phe Asp Phe Asp Leu	465	470	475
Ala Asp Gly Gln Leu Pro Glu Ser Leu Asn Met Glu Glu Ala Tyr Gly	485	490	495
Leu Thr Leu Gln Arg Ala Asp Pro Leu Val Val His Pro Lys Pro Arg	500	505	510

- 132 -

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 971 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..811

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAT ATG CTT AGC ACT TTA ATC TCC CTT AAA GGA ACT GAT CTT GAC GGT 48
Asp Met Leu Ser Thr Leu Ile Ser Leu Lys Gly Thr Asp Leu Asp Gly
1 5 10 15

GAC GGA GGA AGC TTA ACG GAT ACT GAG ATT AAA GCC TTG CTA TTG AAC 96
Asp Gly Gly Ser Leu Thr Asp Thr Glu Ile Lys Ala Leu Leu Leu Asn
20 25 30

ATG TTC ACA GCT GGA ACT GAC ACG TCA GCA AGT ACG GTG GAC TGG GCT 144
Met Phe Thr Ala Gly Thr Asp Thr Ser Ala Ser Thr Val Asp Trp Ala
35 40 45

ATA GCT GAA CTT ATC CGT CAC CCG GAT ATA ATG GTT AAA GCC CAA GAA 192
Ile Ala Glu Leu Ile Arg His Pro Asp Ile Met Val Lys Ala Gln Glu
50 55 60

GAA CTT GAT ATT GTT GTG GGC CGT GAC AGG CCT GTT AAT GAA TCA GAC 240
Glu Leu Asp Ile Val Val Gly Arg Asp Arg Pro Val Asn Glu Ser Asp
65 70 75 80

- 133 -

ATC GCT CAG CTT CCT TAC CTT CAG GCG GTT ATC AAA GAG AAT TTC AGG	288
Ile Ala Gln Leu Pro Tyr Leu Gln Ala Val Ile Lys Glu Asn Phe Arg	
85 90 95	
CTT CAT CCA CCA ACA CCA CTC TCG TTA CCA CAC ATC GCG TCA GAG AGC	336
Leu His Pro Pro Thr Pro Leu Ser Leu Pro His Ile Ala Ser Glu Ser	
100 105 110	
TGT GAG ATC AAC GGC TAC CAT ATC CCG AAA GGA TCG ACT CTA TTT GAC	384
Cys Glu Ile Asn Gly Tyr His Ile Pro Lys Gly Ser Thr Leu Phe Asp	
115 120 125	
GGA CAT ATG GGC CTA GGC CGT GAC CCG GAT CAA TGG TCC GAC CCG TTA	432
Gly His Met Gly Leu Gly Arg Asp Pro Asp Gln Trp Ser Asp Pro Leu	
130 135 140	
GCA TTT AAA CCC GAG AGA TTC TTA CCC GGT GGT GAA AAA TCC GGC GTT	480
Ala Phe Lys Pro Glu Arg Phe Leu Pro Gly Gly Glu Lys Ser Gly Val	
145 150 155 160	
GAT GTG AAA GGA AGC GAT TTC GAG CTA ATA CCG TTC GGG GCT GGG AGG	528
Asp Val Lys Gly Ser Asp Phe Glu Leu Ile Pro Phe Gly Ala Gly Arg	
165 170 175	
CCA ATC TGT GCA GGT TTA AGT TTA GGG CTA CGT ACA GAT TTA AGT TGC	576
Pro Ile Cys Ala Gly Leu Ser Leu Gly Leu Arg Thr Asp Leu Ser Cys	
180 185 190	
CTT CAC GCC AAC GTT GCT CAC AAG CAT TTG ATT GGG AAC TTC AGC TGG	624
Leu His Ala Asn Val Ala His Lys His Leu Ile Gly Asn Phe Ser Trp	
195 200 205	
AGA AGT TAC GCC GGA CAA CCT GAA TAT CGC AGG AAA AGT TTA CTG GGC	672
Arg Ser Tyr Ala Gly Gln Pro Glu Tyr Arg Arg Lys Ser Leu Leu Gly	
210 215 220	
TTT ACA CTG CAA AGA GCG GTT CCT TCG GTG GTA CAC CCT AAG CCA AGG	720

- 134 -

Phe	Thr	Leu	Gln	Arg	Ala	Val	Pro	Ser	Val	Val	His	Pro	Lys	Pro	Arg	
225					230					235					240	
TTG	GCC	CCG	AAC	GTT	TAT	GGA	CCC	CGG	GTC	GGC	TTA	AAA	TTT	AAC	TTT	768
Leu	Ala	Pro	Asn	Val	Tyr	Gly	Pro	Arg	Val	Gly	Leu	Lys	Phe	Asn	Phe	
			245						250				255			
GCT	TCT	TGG	ACA	AGG	TAT	ATG	GCT	TGC	ACG	AAA	CTA	ACG	TTT	T		811
Ala	Ser	Trp	Thr	Arg	Tyr	Met	Ala	Cys	Thr	Lys	Leu	Thr	Phe			
			260						265				270			
AACACACCGT	AGTTTGATCC	GGAGTTAGCT	TTATGTAAGA	ACGTGTAACG	CCAAATCAAG											871
CCATTATCAA	CTACCGTGAG	CTGTTTGTAC	CCTATCTATA	AATCTTGAAG	AGGAACATT											931
CAGAACTCTT	GACTATGTTT	CAGGAACAAA	AAAAAAAAAA													971

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 270 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Asp	Met	Leu	Ser	Thr	Leu	Ile	Ser	Leu	Lys	Gly	Thr	Asp	Leu	Asp	Gly	
1				5					10					15		
Asp	Gly	Gly	Ser	Leu	Thr	Asp	Thr	Glu	Ile	Lys	Ala	Leu	Leu	Leu	Asn	
			20					25				30				
Met	Phe	Thr	Ala	Gly	Thr	Asp	Thr	Ser	Ala	Ser	Thr	Val	Asp	Trp	Ala	
			35					40					45			

- 135 -

Ile Ala Glu Leu Ile Arg His Pro Asp Ile Met Val Lys Ala Gln Glu

50

55

60

Glu Leu Asp Ile Val Val Gly Arg Asp Arg Pro Val Asn Glu Ser Asp

65

70

75

80

Ile Ala Gln Leu Pro Tyr Leu Gln Ala Val Ile Lys Glu Asn Phe Arg

85

90

95

Leu His Pro Pro Thr Pro Leu Ser Leu Pro His Ile Ala Ser Glu Ser

100

105

110

Cys Glu Ile Asn Gly Tyr His Ile Pro Lys Gly Ser Thr Leu Phe Asp

115

120

125

Gly His Met Gly Leu Gly Arg Asp Pro Asp Gln Trp Ser Asp Pro Leu

130

135

140

Ala Phe Lys Pro Glu Arg Phe Leu Pro Gly Gly Glu Lys Ser Gly Val

145

150

155

160

Asp Val Lys Gly Ser Asp Phe Glu Leu Ile Pro Phe Gly Ala Gly Arg

165

170

175

Pro Ile Cys Ala Gly Leu Ser Leu Gly Leu Arg Thr Asp Leu Ser Cys

180

185

190

Leu His Ala Asn Val Ala His Lys His Leu Ile Gly Asn Phe Ser Trp

195

200

205

Arg Ser Tyr Ala Gly Gln Pro Glu Tyr Arg Arg Lys Ser Leu Leu Gly

210

215

220

Phe Thr Leu Gln Arg Ala Val Pro Ser Val Val His Pro Lys Pro Arg

225

230

235

240

Leu Ala Pro Asn Val Tyr Gly Pro Arg Val Gly Leu Lys Phe Asn Phe

- 136 -

245	250	255
Ala Ser Trp Thr Arg Tyr Met Ala Cys Thr Lys Leu Thr Phe		
260	265	270

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6595 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1478..1927

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2651..3091

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3170..3340

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3421..3900

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GTCGACTCTC TCCCTTTCGC TTGCTACTTT TTCTACATAA ATAAATGCAA TGATAAATTT

- 137 -

GTGCACACAT TCGTATGTTT GAAACATGGT AGGATCCACA ATTTATACTT TATAGACTCA	120
AAATGGAAAA GAAACGTACA TTATAAATTT ATCTGCAATT TGTCTTCTCT TGCTAAACTA	180
GACTGTATAA TAACCTCTGT ATATGCTATT ACTCGATTGT AAACGTACCC CGCAAGTCGC	240
AAGCAAGGTA AATAAAGTAT AATTATATTT TCACACACGA AACTTTAATT ATTATTTTTA	300
TCACTTGCAG ATTAACAGTA AAAAAAAAAA AAATGTGACT TTAACGGCGA CAAAAACTAC	360
TGATCTTTCT CCAATATTTA AATAATATAA TTAATAAACG TCTTTTCATA CTTGTATTTT	420
CCGACCCGAG TTCTGAAAGT GAAAACATAT GGTACTAGAT ATTCTCGATT TGTTTTGTAG	480
CCACTAGACT CTAAACAGAA AAAAGAAGCC AAAAGGACAA CGTTAAAAAA GAGACACTGT	540
TATTAAAAGT TAGAAACCAA ACGGTGAAAA TCCAGCTACA TACATAAAAT AAAGCCAAGG	600
TACCAAATA ATGAACTGTA ACCTCTTTTT TCTTTTCTTT TTTGTAAAG GATTTATGAA	660
CTGTAACTTA GAATGCTTGG TTTGTGGGCA GTGTAATATA TGACACACAT GCATTTTTTT	720
TGTTTGTCAG ATAGGAAGAC TTCTTTTTTC TTTATCAACT TCCTTATTTT CATAAAACAA	780
AACACTGAAA AAAGTACAGA TGTCTCACG TACGTCACGT GTACATACAT ATATATTAGA	840
CCACTATATA ATAAGATATG AAGTGTTAGG TTTAAATCAA TTAACGAATC CCATCCAAAT	900
GATGAAACAG TTAACAAGAA ATCAAAATAG TTTATTAGGG TTACAATGAT TTTATACTTT	960
TAAGAAATCT TAGAACCTAT CACTTACAAA TGAGTAAATG ACCATTACTC CTCGAGAATC	1020
TAAGGCGCTT AAGGAAGCAT TGCGAATCGG GTGTGAAAAA GATCTATTTT TTGAATTATT	1080
TCACACAATT TCTTAATGTC AATTTTCGAT GCTCCCATAT TCTCCACGGT TTAAAGCAAG	1140
ATTGGTGGA AAGGGATATT CTCGCATCGA TTACAATGAA ATATGGGTTG AAAAAAAAAA	1200

- 138 -

AAAAAATTA CTCAATGTTG CACCAAAAAC CAGAAAATC TAAGTTGCGC TAATAAAAAA	1260
AAAAGTTATA AACCCAACAT CAAACCAAAA CCGTACTAAA CTGTCCCATA TGAGATTTAG	1320
CTTTAAATAA ATTAGTACTT CTCATAACGA TAACTAAATT AAATTTCCCT AGCCAAGACA	1380
TACATATAGT TTTGATTGAC AAAAAAAAAA AAAACTCCTC TATTATAGC TTGTGTTTTG	1440
TTTCCTCATT TTTCACCTAC CATTCAAACC CAACACT ATG GCA ACT CTA TTT CTC	1495
Met Ala Thr Leu Phe Leu	
1 5	
ACA ATC CTC CTA GCC ACT GTC CTC TTC CTC ATC CTC CGT ATC TTC TCT	1543
Thr Ile Leu Leu Ala Thr Val Leu Phe Leu Ile Leu Arg Ile Phe Ser	
10 15 20	
CAC CGT CGC AAC CGC AGC CAC AAC AAC CGT CTT CCA CCG GGG CCA AAC	1591
His Arg Arg Asn Arg Ser His Asn Asn Arg Leu Pro Pro Gly Pro Asn	
25 30 35	
CCA TGG CCC ATC ATC GGA AAC CTC CCT CAC ATG GGC ACT AAG CCT CAT	1639
Pro Trp Pro Ile Ile Gly Asn Leu Pro His Met Gly Thr Lys Pro His	
40 45 50	
CGA ACC CTT TCC GCC ATG GTT ACT ACT TAC GGC CCT ATC CTC CAC CTC	1687
Arg Thr Leu Ser Ala Met Val Thr Thr Tyr Gly Pro Ile Leu His Leu	
55 60 65 70	
CGA CTA GGG TTC GTA GAC GTC GTG GTC GCC GCT TCT AAA TCC GTG GCC	1735
Arg Leu Gly Phe Val Asp Val Val Val Ala Ala Ser Lys Ser Val Ala	
75 80 85	
GAG CAG TTC TTG AAA ATA CAC GAC GCC AAT TTC GCT AGC CGA CCA CCA	1783
Glu Gln Phe Leu Lys Ile His Asp Ala Asn Phe Ala Ser Arg Pro Pro	
90 95 100	
AAC TCA GGA GCC AAA CAC ATG GCA TAT AAC TAT CAA GAT CTT GTC TTT	1831

- 139 -

Asn	Ser	Gly	Ala	Lys	His	Met	Ala	Tyr	Asn	Tyr	Gln	Asp	Leu	Val	Phe		
105					110					115							
GCA	CCT	TAC	GGA	CAC	CGA	TGG	AGA	CTG	TTG	AGA	AAG	ATT	AGT	TCT	GTT	1879	
Ala	Pro	Tyr	Gly	His	Arg	Trp	Arg	Leu	Leu	Arg	Lys	Ile	Ser	Ser	Val		
120					125					130							
CAT	CTA	TTT	TCA	GCT	AAA	GCT	CTC	GAA	GAT	TTC	AAA	CAT	GTT	CGA	CAG	1927	
His	Leu	Phe	Ser	Ala	Lys	Ala	Leu	Glu	Asp	Phe	Lys	His	Val	Arg	Gln		
135					140					145					150		
GTAAACAAT TATAACGGT ATTCTCATTT TCTAACGCTA TAGCTCACTG GCCTGTAATC																1987	
ATGTCATTTT AATGTTTTGA CTTTTTCTTT ATATATACAT AATTATAATT TATAATTGGG																2047	
ATTTCAAACC CTATCTCTCA CTATTTCAAG ACTAGACCGG ATTGGAATTT GAACTTTTGT																2107	
AATGAATATT AGTATCTGCA CATAAATTTT ATGTTAAAGT TGGGTTTTCT TAAAGTGAAT																2167	
TTATATATTA AAAATATATA AACGATTGGG TTTTACTCAA ATGAATTTAC ATAAGAGCTA																2227	
GGTATAAGTG CAAATATGCA ATACTGTCAT TGTCGTGGAT GTATAAAAGT ATGATCTAAC																2287	
TTTGATGATG CCATGGAAAA ATTGGAAAGT TCAGATCCAG AGGAAACATT GCTTGAATTA																2347	
TAAAATGTAT GGACCACATT GTTTCCTTAA ATGGAAGGTC TCACGAGTTT CTCAATTTC																2407	
GACTACTGAT AATATATGCT ATTATAGATT TTATTTTCTG ATTATTTTTT TTGGTTTAAT																2467	
TTAATTAGAG TAAATTTTTT AAAAGAAATA TATGGTTTTG TTAACCGTGT TTTAAAATTT																2527	
GATAGAGCTT TTAGATCATA ATCATAATTT TTTCGTATTA ATTGTGATTA TGTGACGA																2587	
AAATACTTAA TTAGTATTCA AGAAAACTCT TATTCTAAAA ACAGAAATAA ATGAATTTTA																2647	
CAG GAA GAG GTT GGA ACG CTA ACG CGG GAG CTA GTG CGT GTT GGC ACG																2695	
Glu Glu Val Gly Thr Leu Thr Arg Glu Leu Val Arg Val Gly Thr																	

- 140 -

1	5	10	15	
AAA CCC GTG AAT TTA GGC CAG TTG GTG AAC ATG TGT GTA GTC AAC GCT				2743
Lys Pro Val Asn Leu Gly Gln Leu Val Asn Met Cys Val Val Asn Ala				
20	25	30		
CTA GGA CGA GAG ATG ATC GGA CGG CGA CTG TTC GGC GCC GAC GCC GAT				2791
Leu Gly Arg Glu Met Ile Gly Arg Arg Leu Phe Gly Ala Asp Ala Asp				
35	40	45		
CAT AAA GCT GAC GAG TTT CGA TCG ATG GTG ACG GAA ATG ATG GCT CTC				2839
His Lys Ala Asp Glu Phe Arg Ser Met Val Thr Glu Met Met Ala Leu				
50	55	60		
GCC GGA GTA TTT AAC ATC GGA GAT TTC GTG CCG TCA CTT GAT TGG TTA				2887
Ala Gly Val Phe Asn Ile Gly Asp Phe Val Pro Ser Leu Asp Trp Leu				
65	70	75		
GAT TTA CAA GGC GTC GCT GGT AAA ATG AAA CGG CTT CAC AAA AGA TTC				2935
Asp Leu Gln Gly Val Ala Gly Lys Met Lys Arg Leu His Lys Arg Phe				
80	85	90	95	
GAC GCT TTT CTA TCG TCG ATT TTG AAA GAG CAC GAA ATG AAC GGT CAA				2983
Asp Ala Phe Leu Ser Ser Ile Leu Lys Glu His Glu Met Asn Gly Gln				
100	105	110		
GAT CAA AAG CAT ACA GAT ATG CTT AGC ACT TTA ATC TCC CTT AAA GGA				3031
Asp Gln Lys His Thr Asp Met Leu Ser Thr Leu Ile Ser Leu Lys Gly				
115	120	125		
ACT GAT CTT GAC GGT GAC GGA GGA AGC TTA ACG GAT ACT GAG ATT AAA				3079
Thr Asp Leu Asp Gly Asp Gly Gly Ser Leu Thr Asp Thr Glu Ile Lys				
130	135	140		
GCC TTG CTA TTG GTCAGTTTTT TGACAATTAA TTTCCTTAAA AATCGTATAT				3131
Ala Leu Leu Leu				
145				

- 141 -

AATGAAAGTT AGATTGTTTT TTTTGTTGT AAATACAG AAC ATG TTC ACA GCT 3184

Asn Met Phe Thr Ala

1

5

GGA ACT GAC ACG TCA GCA AGT ACG GTG GAC TGG GCT ATA GCT GAA CTT 3232

Gly Thr Asp Thr Ser Ala Ser Thr Val Asp Trp Ala Ile Ala Glu Leu

10

15

20

ATC CGT CAC CCG GAT ATA ATG GTT AAA GCC CAA GAA GAA CTT GAT ATT 3280

Ile Arg His Pro Asp Ile Met Val Lys Ala Gln Glu Glu Leu Asp Ile

25

30

35

GTT GTG GGC CGT GAC AGG CCT GTT AAT GAA TCA GAC ATC GCT CAG CTT 3328

Val Val Gly Arg Asp Arg Pro Val Asn Glu Ser Asp Ile Ala Gln Leu

40

45

50

CCT TAC CTT CAG GTACCGTTAA CCCAAACCGG AATTTGGAAT TGTTTTGGTT 3380

Pro Tyr Leu Gln

55

AGCGAGCTAT TGTTGTTAAT CCGGTTTTGG TTAAACAG GCG GTT ATC AAA GAG 3435

Ala Val Ile Lys Glu

1

5

AAT TTC AGG CTT CAT CCA CCA ACA CCA CTC TCG TTA CCA CAC ATC GCG 3483

Asn Phe Arg Leu His Pro Pro Thr Pro Leu Ser Leu Pro His Ile Ala

10

15

20

TCA GAG AGC TGT GAG ATC AAC GGC TAC CAT ATC CCG AAA GGA TCG ACT 3531

Ser Glu Ser Cys Glu Ile Asn Gly Tyr His Ile Pro Lys Gly Ser Thr

25

30

35

CTA TTG ACG AAC ATA TGG GCC ATA GCC CGT GAC CCG GAT CAA TGG TCC 3579

Leu Leu Thr Asn Ile Trp Ala Ile Ala Arg Asp Pro Asp Gln Trp Ser

40

45

50

GAC CCG TTA GCA TTT AAA CCC GAG AGA TTC TTA CCC GGT GGT GAA AAA 3627

- 142 -

[illegible]

- 143 -

TATATACTAG	GTGGAACACC	ACCACCTGCA	ACTCTGCAAC	ACATGTTACG	TTACACAATC	4277
ACTTTTGGCG	GTTTTCAATT	ATTTATATAA	AATTGTAAAT	GTTTGTACAC	AGTAGAAAAT	4337
TAGTAATAGT	GAATTTTGTT	TCTCCGAATA	TGTATAGCAA	TATATATGGC	ATGGATCAAA	4397
CTAGCCGACA	TCCTAACTTG	TTCACAGCTT	TCCTTTTTAC	TTATCTAGTC	GATTAAGCAT	4457
CAGAAAGTAT	GTTTTAATTT	TTAAATTTGA	AAAAGGTGTA	CTTACAAGTT	CGGGTGTTC	4517
CACGGAGGAG	AGCTACAATA	ATGAAAAGC	TGACTCAAGA	AGGGCTATAG	AAGAAACAAG	4577
AGTCACGGAA	CAAGTTGTCA	CTCTCAATCT	CCAGTACACT	AGCTTCCATA	ACTCTCTCTC	4637
TTTCTCTCTT	TCTTCTCTCT	CTAAAAGTTA	TCAGAATAGA	AATCTCTCTC	TCTCAACAAG	4697
TCTAACAGTG	CCATTTGTAT	CTCTGAACTC	CAACATGGCT	CCTCTGGTTC	TCTACCTTCT	4757
CACTCTCCTC	ATGGCTGGCC	ATTCCAGTAA	GAACTCTCAC	TGATCTTCTT	CACCTTTGTT	4817
TATGGATTTG	GTCTCTCAGT	CTCACTCTCG	CTTACCCTTT	CACATTCAGC	TCTGGCTCTC	4877
TGGTTTAAGA	AACCCTTAAT	CTACAAAGCT	TGCTTTCCTC	GCAAATGAAC	TACCTTACTT	4937
ATCTCTTATG	CAACTCTTGT	TGATGATTTG	CAAACATCTT	AACCTCTCGA	AACAAGATTT	4997
ACAAATCTTA	CTGGCTTCAC	TTACAATTTT	GTTCCCATTT	TTTTCTTCTT	TGGTAGGTGC	5057
CTCATGGTGT	GTGTGCAAAA	CAGGGCTGAG	TGACTCAGTG	CTACAAAAGA	CATTAGACTA	5117
TGCTTGTGGA	AATGGAGCTG	ACTGTAACCC	AACTCACCCA	AAAGGCTCTT	GCTTCAATCC	5177
TGACAATGTT	AGGGCTCATT	GCAACTATGC	AGTCAATAGC	TTCTTCCAAA	AGAAAGGTCA	5237
AGCTTCTGAG	TCTTGTAAC	TCACTGGTAC	TGCCACTCTT	ACCACCACCG	ATCCCAGTAA	5297
GTTTTTCAGAA	TGTTAACACT	CTTGTGATCT	TTAGAACCCT	ACAAAATTTT	GAGTCTCAGA	5357

- 144 -

AAGTTCAAGT TCAAGGTCTT TTGGTTAGAG TACTAAAGAT TCAAGTAGAG ACTAGGCGTG 5417

AGATATTTTT TCTCTGATGT GTGATTTTTT GGCACAGGCT ATACAGGATG TGCATTCCCT 5477

TCTAGTGCTA GGTACGGCTC TTTGCTTCTC TACACATTTA TTTTCTTAAT GGCTTTATCT 5537

AGAACTTTGA AGGATACCAT TTTATTTTTT TTGGACAAAG AAGGATAGCC ATTTAATACT 5597

ACACTTTAAT GTTGGATTAA CTAAC TTATT ATGCCTATTT AATGGCCTAC ACTTTAAGTG 5657

GACACAAGCT TGATTTGGTT ATAAAAAAG TGCATAATA TCTTATTTTA CTGAACCCTT 5717

TTTTCTATGA TTTTTTTACT AAAC TTAGA TAACATCTAC AACAAATCAA TTGCCTTTTT 5777

TTGGGGATTG TATAAGTTTG AACCTATGGT TAGTGTATTG ACTTGCGCGT CTCTTATTGC 5837

AACGGTTCTT TGAAAACACA TTAATGATAA ATAAATTGAA AAGTATAGAG ATGGCAATTG 5897

TTTCAAAGC TAATCTTTCT GCTTGCTAAT ACTTTACATA AAAACAAAA AATTAAGAAG 5957

ATTTTCAAAC AATACAAC TT TTTACCTTG TCCTAACAAA TTCAACTCAA ATGACATGTG 6017

TTTGCTTTAA AATAGTAACA ACTGTAAATT CATTTGCTCT TGAGACATAA GTGCAAGCTA 6077

AAGATAAACG CAAGCAATAC AATTAGGCCT AATTAAGATT ACGAATATTG TTGTTTGTTT 6137

ATAGTGGTTC TAGTGGAAGC GGTAGCACCA CCGTGACGCC AGGCAAAAAC AGTCCAAAAG 6197

GAAGCAACAG CATCACCACA TTTCCCGGCG GAAACAGTCC ATACACTGGC ACACCATCCA 6257

CCGGATTATT AGGAGGCAAT ATCACTGATG CAACTGGAAC CGGGTTGAAC CCGGATTACT 6317

CAACCGAAAG CAGTGGATTT GCGCTCTATT ACTCCAACAA CCTTCTGTTA ACCGGCTTTT 6377

GTTCTCTCGT GATGATGCTC TGAAGAAGAA TCACCGTCTT CTTT TAGTTT ATGCTTAGTC 6437

AAAAAATAT GTTATTTATA TGTCTTGTT GTTTTAGAGA TAATTTAATC TGGATTTTCGG 6497

- 145 -

TTCTTTTTTA CTTTCGGTT TTAAGAAAAC AATTATCAAT GTAAAACCAA ATCTACTATC 6557

GATCGGTTTG GTACGAATTC CTGCAGCCCCG GGGGATCC 6595

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 149 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Ala Thr Leu Phe Leu Thr Ile Leu Leu Ala Thr Val Leu Phe Leu
1 5 10 15

Ile Leu Arg Ile Phe Ser His Arg Arg Asn Arg Ser His Asn Asn Arg
20 25 30

Leu Pro Pro Gly Pro Asn Pro Trp Pro Ile Ile Gly Asn Leu Pro His
35 40 45

Met Gly Thr Lys Pro His Arg Thr Leu Ser Ala Met Val Thr Thr Tyr
50 55 60

Gly Pro Ile Leu His Leu Arg Leu Gly Phe Val Asp Val Val Val Ala
65 70 75 80

Ala Ser Lys Ser Val Ala Glu Gln Phe Leu Lys Ile His Asp Ala Asn
85 90 95

Phe Ala Ser Arg Pro Pro Asn Ser Gly Ala Lys His Met Ala Tyr Asn
100 105 110

- 146 -

Tyr Gln Asp Leu Val Phe Ala Pro Tyr Gly His Arg Trp Arg Leu Leu

115

120

125

Arg Lys Ile Ser Ser Val His Leu Phe Ser Ala Lys Ala Leu Glu Asp

130

135

140

Phe Lys His Val Arg Gln

145

150

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 147 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Glu Glu Val Gly Thr Leu Thr Arg Glu Leu Val Arg Val Gly Thr Lys

1

5

10

15

Pro Val Asn Leu Gly Gln Leu Val Asn Met Cys Val Val Asn Ala Leu

20

25

30

Gly Arg Glu Met Ile Gly Arg Arg Leu Phe Gly Ala Asp Ala Asp His

35

40

45

Lys Ala Asp Glu Phe Arg Ser Met Val Thr Glu Met Met Ala Leu Ala

50

55

60

Gly Val Phe Asn Ile Gly Asp Phe Val Pro Ser Leu Asp Trp Leu Asp

65

70

75

80

Leu Gln Gly Val Ala Gly Lys Met Lys Arg Leu His Lys Arg Phe Asp

- 147 -

	85	90	95
Ala Phe Leu Ser Ser Ile Leu Lys Glu His Glu Met Asn Gly Gln Asp			
100	105	110	
Gln Lys His Thr Asp Met Leu Ser Thr Leu Ile Ser Leu Lys Gly Thr			
115	120	125	
Asp Leu Asp Gly Asp Gly Gly Ser Leu Thr Asp Thr Glu Ile Lys Ala			
130	135	140	
Leu Leu Leu			
145			

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 57 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Asn Met Phe Thr Ala Gly Thr Asp Thr Ser Ala Ser Thr Val Asp Trp			
1	5	10	15
Ala Ile Ala Glu Leu Ile Arg His Pro Asp Ile Met Val Lys Ala Gln			
20	25	30	
Glu Glu Leu Asp Ile Val Val Gly Arg Asp Arg Pro Val Asn Glu Ser			
35	40	45	
Asp Ile Ala Gln Leu Pro Tyr Leu Gln			
50	55		

- 148 -

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 160 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Ala	Val	Ile	Lys	Glu	Asn	Phe	Arg	Leu	His	Pro	Pro	Thr	Pro	Leu	Ser
1				5					10					15	
Leu	Pro	His	Ile	Ala	Ser	Glu	Ser	Cys	Glu	Ile	Asn	Gly	Tyr	His	Ile
			20					25					30		
Pro	Lys	Gly	Ser	Thr	Leu	Leu	Thr	Asn	Ile	Trp	Ala	Ile	Ala	Arg	Asp
		35					40						45		
Pro	Asp	Gln	Trp	Ser	Asp	Pro	Leu	Ala	Phe	Lys	Pro	Glu	Arg	Phe	Leu
	50					55					60				
Pro	Gly	Gly	Glu	Lys	Ser	Gly	Val	Asp	Val	Lys	Gly	Ser	Asp	Phe	Glu
65				70						75				80	
Leu	Ile	Pro	Phe	Gly	Ala	Gly	Arg	Arg	Ile	Cys	Ala	Gly	Leu	Ser	Leu
			85						90					95	
Gly	Leu	Arg	Thr	Ile	Gln	Phe	Leu	Thr	Ala	Thr	Leu	Val	Gln	Gly	Phe
			100						105					110	
Asp	Trp	Glu	Leu	Ala	Gly	Gly	Val	Thr	Pro	Glu	Lys	Leu	Asn	Met	Glu
		115						120						125	
Glu	Ser	Tyr	Gly	Leu	Thr	Leu	Gln	Arg	Ala	Val	Pro	Leu	Val	Val	His

- 149 -

130	135	140
Pro Lys Pro Arg Leu Ala Pro Asn Val Tyr Gly Leu Gly Ser Gly		
145	150	155

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1748 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 22..1563

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TGTCGAGAAA GAAGAACAGC C ATG TTT CTC ATA GTA GTG ATC ACC TTC CTC	51
Met Phe Leu Ile Val Val Ile Thr Phe Leu	
1 5 10	
TTC GCC GTG TTT TTG TTC CGG CTT CTT TTC TCC GGC AAA TCC CAA CGC	99
Phe Ala Val Phe Leu Phe Arg Leu Leu Phe Ser Gly Lys Ser Gln Arg	
15 20 25	
CAC TCG CTC CCT CTC CCT CCT GGC CCC AAA CCA TGG CCG GTG GTT GGC	147
His Ser Leu Pro Leu Pro Pro Gly Pro Lys Pro Trp Pro Val Val Gly	
30 35 40	

- 150 -

AAC TTA CCT CAC TTG GGC CCC TTC CCG CAC CAC TCC ATC GCG GAG TTG	195
Asn Leu Pro His Leu Gly Pro Phe Pro His His Ser Ile Ala Glu Leu	
45 50 55	
GCG AAG AAA CAC GGG CCG CTC ATG CAC CTC CGC CTC GGC TAC GTT GAC	243
Ala Lys Lys His Gly Pro Leu Met His Leu Arg Leu Gly Tyr Val Asp	
60 65 70	
GTA GTC GTG GCG GCA TCA GCA TCC GTA GCG GCC CAG TTC TTG AAG ACT	291
Val Val Val Ala Ala Ser Ala Ser Val Ala Ala Gln Phe Leu Lys Thr	
75 80 85 90	
CAC GAC GCC AAT TTC TCC AGC CGA CCG CCC AAC TCC GGC GCC AAG CAC	339
His Asp Ala Asn Phe Ser Ser Arg Pro Pro Asn Ser Gly Ala Lys His	
95 100 105	
CTC GCC TAT AAC TAC CAG GAC CTC GTG TTC AGG CCG TAC GGT CCA CGG	387
Leu Ala Tyr Asn Tyr Gln Asp Leu Val Phe Arg Pro Tyr Gly Pro Arg	
110 115 120	
TGG CGC ATG TTC CGG AAG ATC AGC TCC GTC CAT CTG TTC TCC GGC AAA	435
Trp Arg Met Phe Arg Lys Ile Ser Ser Val His Leu Phe Ser Gly Lys	
125 130 135	
GCC TTG GAT GAT CTT AAA CAC GTC CGG CAG GAG GAG GTA AGT GTG CTA	483
Ala Leu Asp Asp Leu Lys His Val Arg Gln Glu Glu Val Ser Val Leu	
140 145 150	
GCG CAT GCC TTG GCA AAT TCA GGG TCA AAG GTA GTG AAC CTG GCG CAA	531
Ala His Ala Leu Ala Asn Ser Gly Ser Lys Val Val Asn Leu Ala Gln	
155 160 165 170	
CTG CTG AAC CTG TGC ACG GTC AAT GCT CTA GGA AGG GTG ATG GTA GGG	579
Leu Leu Asn Leu Cys Thr Val Asn Ala Leu Gly Arg Val Met Val Gly	
175 180 185	
CGG AGG GTT TTC GGC GAC GGC AGC GGA GGC GAC GAT CCG AAG GCG GAC	627

- 151 -

Arg	Arg	Val	Phe	Gly	Asp	Gly	Ser	Gly	Gly	Asp	Asp	Pro	Lys	Ala	Asp		
				190					195					200			
GAG	TTC	AAA	TCG	ATG	GTG	GTG	GAG	ATG	ATG	GTG	TTG	GCA	GGA	GTG	TTC	675	
Glu	Phe	Lys	Ser	Met	Val	Val	Glu	Met	Met	Val	Leu	Ala	Gly	Val	Phe		
				205					210					215			
AAC	ATA	GGT	GAC	TTC	ATC	CCC	TCT	CTC	GAA	TGG	CTT	GAC	TTG	CAA	GGC	723	
Asn	Ile	Gly	Asp	Phe	Ile	Pro	Ser	Leu	Glu	Trp	Leu	Asp	Leu	Gln	Gly		
				220					225					230			
GTG	GCG	TCC	AAG	ATG	AAG	AAG	CTC	CAC	AAG	AGA	TTC	GAC	GAC	TTC	TTG	771	
Val	Ala	Ser	Lys	Met	Lys	Lys	Leu	His	Lys	Arg	Phe	Asp	Asp	Phe	Leu		
				235					240					245			250
ACA	GCC	ATT	GTC	GAG	GAC	CAC	AAG	AAG	GGC	TCC	GGC	ACG	GCG	GGG	CAC	819	
Thr	Ala	Ile	Val	Glu	Asp	His	Lys	Lys	Gly	Ser	Gly	Thr	Ala	Gly	His		
				255					260					265			
GTC	GAC	ATG	TTG	ACC	ACT	CTG	CTC	TCG	CTC	AAG	GAA	GAC	GCC	GAC	GGC	867	
Val	Asp	Met	Leu	Thr	Thr	Leu	Leu	Ser	Leu	Lys	Glu	Asp	Ala	Asp	Gly		
				270					275					280			
GAA	GGA	GGC	AAG	CTC	ACC	GAT	ACT	GAA	ATC	AAA	GCT	TTG	CTT	TTG	AAC	915	
Glu	Gly	Gly	Lys	Leu	Thr	Asp	Thr	Glu	Ile	Lys	Ala	Leu	Leu	Leu	Asn		
				285					290					295			
ATG	TTC	ACG	GCT	GGC	ACT	GAT	ACG	TCA	TCG	AGC	ACG	GTG	GAA	TGG	GCA	963	
Met	Phe	Thr	Ala	Gly	Thr	Asp	Thr	Ser	Ser	Ser	Thr	Val	Glu	Trp	Ala		
				300					305					310			
ATA	GCT	GAA	CTC	ATT	CGG	CAC	CCT	CAT	ATG	CTA	GCG	CGA	GTT	CAG	AAA	1011	
Ile	Ala	Glu	Leu	Ile	Arg	His	Pro	His	Met	Leu	Ala	Arg	Val	Gln	Lys		
				315					320					325			330
GAG	CTT	GAC	GAT	TTT	GTT	GGC	CAT	GAC	CGA	CTT	GTG	ACC	GAA	TCC	GAC	1059	
Glu	Leu	Asp	Asp	Phe	Val	Gly	His	Asp	Arg	Leu	Val	Thr	Glu	Ser	Asp		

- 152 -

	335	340	345	
ATA CCC AAC CTC CCT TAC CTC CAA GCC GTG ATC AAG GAA ACG TTC CGA				1107
Ile Pro Asn Leu Pro Tyr Leu Gln Ala Val Ile Lys Glu Thr Phe Arg				
	350	355	360	
CTC CAC CCA TCC ACT CCT CTC TCG TTG CCT CGT ATG GCA GCC GAG AGT				1155
Leu His Pro Ser Thr Pro Leu Ser Leu Pro Arg Met Ala Ala Glu Ser				
	365	370	375	
TGC GAA ATC AAC GGG TAC CAC ATC CCG AAA GGC TCC ACA CTC TTG GTC				1203
Cys Glu Ile Asn Gly Tyr His Ile Pro Lys Gly Ser Thr Leu Leu Val				
	380	385	390	
AAT GTA TGG GCC ATA TCG CGT GAC CCG GCT GAA TGG GCC GAC CCA CTG				1251
Asn Val Trp Ala Ile Ser Arg Asp Pro Ala Glu Trp Ala Asp Pro Leu				
	395	400	405	410
GAG TTC AAG CCC GAG AGG TTC CTG CCG GGG GGC GAA AAG CCT AAT GTT				1299
Glu Phe Lys Pro Glu Arg Phe Leu Pro Gly Gly Glu Lys Pro Asn Val				
	415	420	425	
GAT ATT AGA GGA AAC GAT TTT GAA GTC ATA CCC TTC GGT GCC GGG CGA				1347
Asp Ile Arg Gly Asn Asp Phe Glu Val Ile Pro Phe Gly Ala Gly Arg				
	430	435	440	
AGA ATA TGT GCC GGG ATG AGC TTG GGC CTG CGT ATG GTC CAT TTA ATG				1395
Arg Ile Cys Ala Gly Met Ser Leu Gly Leu Arg Met Val His Leu Met				
	445	450	455	
ACT GCA ACA TTG GTC CAC GCA TTT AAT TGG GCC TTG GCT GAT GGG CTG				1443
Thr Ala Thr Leu Val His Ala Phe Asn Trp Ala Leu Ala Asp Gly Leu				
	460	465	470	
ACC GCT GAG AAG TTA AAC ATG GAT GAA GCA TAT GGG CTC ACT CTA CAA				1491
Thr Ala Glu Lys Leu Asn Met Asp Glu Ala Tyr Gly Leu Thr Leu Gln				
	475	480	485	490

- 153 -

CGA GCT GCA CCG TTA ATG GTG CAC CCG CGC ACC AGG CTG GCC CCA CAG 1539

Arg Ala Ala Pro Leu Met Val His Pro Arg Thr Arg Leu Ala Pro Gln

495

500

505

GCA TAT AAA ACT TCA TCA TCT TAATTAGAGA GCTATGTTCT GGGTGTGCCC 1590

Ala Tyr Lys Thr Ser Ser Ser

510

GGTTTGATGT CTCCATGTTT TCTATTTAGG TTAAATCTG TAAGATAAGG TGATTCTATG 1650

CTGAATCACA AAAGTTGCTA TCTAAATTCC ATGTCCAATG AAAACGTTCT TCTTCCCTTC 1710

TTATAATTTA TGAATACTTA TGATATAGGC GACAGCAA 1748

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 513 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met Phe Leu Ile Val Val Ile Thr Phe Leu Phe Ala Val Phe Leu Phe

1

5

10

15

Arg Leu Leu Phe Ser Gly Lys Ser Gln Arg His Ser Leu Pro Leu Pro

20

25

30

Pro Gly Pro Lys Pro Trp Pro Val Val Gly Asn Leu Pro His Leu Gly

35

40

45

Pro Phe Pro His His Ser Ile Ala Glu Leu Ala Lys Lys His Gly Pro

50

55

60

- 154 -

Leu Met His Leu Arg Leu Gly Tyr Val Asp Val Val Val Ala Ala Ser

65

70

75

80

Ala Ser Val Ala Ala Gln Phe Leu Lys Thr His Asp Ala Asn Phe Ser

85

90

95

Ser Arg Pro Pro Asn Ser Gly Ala Lys His Leu Ala Tyr Asn Tyr Gln

100

105

110

Asp Leu Val Phe Arg Pro Tyr Gly Pro Arg Trp Arg Met Phe Arg Lys

115

120

125

Ile Ser Ser Val His Leu Phe Ser Gly Lys Ala Leu Asp Asp Leu Lys

130

135

140

His Val Arg Gln Glu Glu Val Ser Val Leu Ala His Ala Leu Ala Asn

145

150

155

160

Ser Gly Ser Lys Val Val Asn Leu Ala Gln Leu Leu Asn Leu Cys Thr

165

170

175

Val Asn Ala Leu Gly Arg Val Met Val Gly Arg Arg Val Phe Gly Asp

180

185

190

Gly Ser Gly Gly Asp Asp Pro Lys Ala Asp Glu Phe Lys Ser Met Val

195

200

205

Val Glu Met Met Val Leu Ala Gly Val Phe Asn Ile Gly Asp Phe Ile

210

215

220

Pro Ser Leu Glu Trp Leu Asp Leu Gln Gly Val Ala Ser Lys Met Lys

225

230

235

240

Lys Leu His Lys Arg Phe Asp Asp Phe Leu Thr Ala Ile Val Glu Asp

245

250

255

His Lys Lys Gly Ser Gly Thr Ala Gly His Val Asp Met Leu Thr Thr

- 155 -

260	265	270
Leu Leu Ser Leu Lys Glu Asp Ala Asp Gly Glu Gly Gly Lys Leu Thr		
275	280	285
Asp Thr Glu Ile Lys Ala Leu Leu Leu Asn Met Phe Thr Ala Gly Thr		
290	295	300
Asp Thr Ser Ser Ser Thr Val Glu Trp Ala Ile Ala Glu Leu Ile Arg		
305	310	315 320
His Pro His Met Leu Ala Arg Val Gln Lys Glu Leu Asp Asp Phe Val		
325	330	335
Gly His Asp Arg Leu Val Thr Glu Ser Asp Ile Pro Asn Leu Pro Tyr		
340	345	350
Leu Gln Ala Val Ile Lys Glu Thr Phe Arg Leu His Pro Ser Thr Pro		
355	360	365
Leu Ser Leu Pro Arg Met Ala Ala Glu Ser Cys Glu Ile Asn Gly Tyr		
370	375	380
His Ile Pro Lys Gly Ser Thr Leu Leu Val Asn Val Trp Ala Ile Ser		
385	390	395 400
Arg Asp Pro Ala Glu Trp Ala Asp Pro Leu Glu Phe Lys Pro Glu Arg		
405	410	415
Phe Leu Pro Gly Gly Glu Lys Pro Asn Val Asp Ile Arg Gly Asn Asp		
420	425	430
Phe Glu Val Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala Gly Met		
435	440	445
Ser Leu Gly Leu Arg Met Val His Leu Met Thr Ala Thr Leu Val His		
450	455	460

- 156 -

Ala Phe Asn Trp Ala Leu Ala Asp Gly Leu Thr Ala Glu Lys Leu Asn

465

470

475

480

Met Asp Glu Ala Tyr Gly Leu Thr Leu Gln Arg Ala Ala Pro Leu Met

485

490

495

Val His Pro Arg Thr Arg Leu Ala Pro Gln Ala Tyr Lys Thr Ser Ser

500

505

510

Ser

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1660 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 4..1528

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

AAA ATG ACC ATT TTA GCT TTC GTA TTT TAC GCC CTC ATC CTC GGG TCA 48

Met Thr Ile Leu Ala Phe Val Phe Tyr Ala Leu Ile Leu Gly Ser

1

5

10

15

GTA CTC TAT GTA TTT CTT AAC TTA AGT TCA CGT AAA TCC GCC AGA CTC 96

- 157 -

Val Leu Tyr Val Phe Leu Asn Leu Ser Ser Arg Lys Ser Ala Arg Leu

20

25

30

CCA CCC GGG CCA ACA CCA TGG CCT ATA GTC GGG AAC TTA CCA CAC CTT 144

Pro Pro Gly Pro Thr Pro Trp Pro Ile Val Gly Asn Leu Pro His Leu

35

40

45

GGC CCA ATC CCA CAC CAC GCA CTC GCG GCC TTA GCC AAG AAG TAC GGG 192

Gly Pro Ile Pro His His Ala Leu Ala Ala Leu Ala Lys Lys Tyr Gly

50

55

60

CCA TTG ATG CAC CTG CGG CTC GGG TGT GTG GAC GTG GTT GTG GCC GCG 240

Pro Leu Met His Leu Arg Leu Gly Cys Val Asp Val Val Val Ala Ala

65

70

75

TCT GCT TCC GTA GCT GCA CAG TTT TTA AAA GTT CAC GAC GCA AAT TTT 288

Ser Ala Ser Val Ala Ala Gln Phe Leu Lys Val His Asp Ala Asn Phe

80

85

90

95

GCT AGT AGG CCG CCA AAT TCT GGC GCG AAA CAT GTG GCG TAT AAT TAT 336

Ala Ser Arg Pro Pro Asn Ser Gly Ala Lys His Val Ala Tyr Asn Tyr

100

105

110

CAG GAT CTT GTG TTT GCA CCT TAT GGT CCA AGG TGG CGT TTG TTA AGG 384

Gln Asp Leu Val Phe Ala Pro Tyr Gly Pro Arg Trp Arg Leu Leu Arg

115

120

125

AAG ATT TGT TCG GTC CAT TTG TTT TCT GCT AAA GCA CTT GAT GAT TTT 432

Lys Ile Cys Ser Val His Leu Phe Ser Ala Lys Ala Leu Asp Asp Phe

130

135

140

CGT CAT GTT CGA CAG GAG GAG GTA GCA GTC CTA ACC CGC GTA CTA CTG 480

Arg His Val Arg Gln Glu Glu Val Ala Val Leu Thr Arg Val Leu Leu

145

150

155

AGT GCT GGA AAC TCA CCG GTA CAG CTT GGC CAA CTA CTT AAC GTG TGT 528

Ser Ala Gly Asn Ser Pro Val Gln Leu Gly Gln Leu Leu Asn Val Cys

- 158 -

160	165	170	175	
GCC ACA AAC GCC TTA GCA CGG GTA ATG TTA GGT AGG AGA GTT TTC GGA				576
Ala Thr Asn Ala Leu Ala Arg Val Met Leu Gly Arg Arg Val Phe Gly				
	180	185	190	
GAC GGA ATT GAC AGG TCA GCC AAT GAG TTC AAA GAT ATG GTA GTA GAG				624
Asp Gly Ile Asp Arg Ser Ala Asn Glu Phe Lys Asp Met Val Val Glu				
	195	200	205	
TTA ATG GTA TTA GCA GGA GAA TTT AAC CTT GGT GAC TTT ATT CCT GTA				672
Leu Met Val Leu Ala Gly Glu Phe Asn Leu Gly Asp Phe Ile Pro Val				
	210	215	220	
CTT GAC CTA TTC GAC CTA CAA GGC ATT ACT AAA AAA ATG AAG AAG CTT				720
Leu Asp Leu Phe Asp Leu Gln Gly Ile Thr Lys Lys Met Lys Lys Leu				
	225	230	235	
CAT GTT CGG TTC GAT TCA TTT CTT AGT AAG ATC GTT GAG GAG CAT AAA				768
His Val Arg Phe Asp Ser Phe Leu Ser Lys Ile Val Glu Glu His Lys				
	240	245	250	255
ACG GCA CCT GGT GGG TTG GGT CAT ACT GAT TTG CTG AGC ACG TTG ATT				816
Thr Ala Pro Gly Gly Leu Gly His Thr Asp Leu Leu Ser Thr Leu Ile				
	260	265	270	
TCA CTT AAA GAT GAT GCT GAT ATT GAA GGT GGG AAG CTT ACA GAT ACT				864
Ser Leu Lys Asp Asp Ala Asp Ile Glu Gly Gly Lys Leu Thr Asp Thr				
	275	280	285	
GAA ATC AAA GCT TTG CTT CTG AAT TTA TTC GCT GCG GGA ACA GAC ACA				912
Glu Ile Lys Ala Leu Leu Leu Asn Leu Phe Ala Ala Gly Thr Asp Thr				
	290	295	300	
TCC TCT AGT ACA GTA GAA TGG GCA ATA GCC GAA CTC ATT CGT CAT CCA				960
Ser Ser Ser Thr Val Glu Trp Ala Ile Ala Glu Leu Ile Arg His Pro				
	305	310	315	

- 159 -

CAA ATA TTA AAA CAA GCC CGA GAA GAG ATA GAC GCT GTA GTT GGT CAA	1008
Gln Ile Leu Lys Gln Ala Arg Glu Glu Ile Asp Ala Val Val Gly Gln	
320 325 330 335	
GAC CGG CTT GTA ACA GAA TTG GAC TTG AGC CAA CTA ACA TAC CTC CAG	1056
Asp Arg Leu Val Thr Glu Leu Asp Leu Ser Gln Leu Thr Tyr Leu Gln	
340 345 350	
GCT CTT GTG AAA GAG GTG TTT AGG CTC CAC CCT TCA ACG CCA CTC TCC	1104
Ala Leu Val Lys Glu Val Phe Arg Leu His Pro Ser Thr Pro Leu Ser	
355 360 365	
TTA CCA AGA ATA TCA TCC GAG AGT TGT GAG GTC GAT GGG TAT TAT ATC	1152
Leu Pro Arg Ile Ser Ser Glu Ser Cys Glu Val Asp Gly Tyr Tyr Ile	
370 375 380	
CCT AAG GGA TCC ACA CTC CTC GTT AAC GTG TGG GCC ATT GCG CGA GAC	1200
Pro Lys Gly Ser Thr Leu Leu Val Asn Val Trp Ala Ile Ala Arg Asp	
385 390 395	
CCA AAA ATG TGG GCG GAT CCT CTT GAA TTT AGG CCT TCT CGG TTT TTA	1248
Pro Lys Met Trp Ala Asp Pro Leu Glu Phe Arg Pro Ser Arg Phe Leu	
400 405 410 415	
CCC GGG GGA GAA AAG CCC GGT GCT GAT GTT AGG GGA AAT GAT TTT GAA	1296
Pro Gly Gly Glu Lys Pro Gly Ala Asp Val Arg Gly Asn Asp Phe Glu	
420 425 430	
GTT ATA CCA TTT GGG GCA GGA CGA AGG ATT TGT GCG GGT ATG AGC CTA	1344
Val Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala Gly Met Ser Leu	
435 440 445	
GGC TTG AGA ATG GTC CAG TTG CTC ATT GCA ACA TTG GTC CAA ACT TTT	1392
Gly Leu Arg Met Val Gln Leu Leu Ile Ala Thr Leu Val Gln Thr Phe	
450 455 460	
GAT TGG GAA CTG GCT AAC GGG TTA GAG CCG GAG ATG CTC AAC ATG GAA	1440

- 160 -

[illegible]

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 508 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```

Met Thr Ile Leu Ala Phe Val Phe Tyr Ala Leu Ile Leu Gly Ser Val
  1                               5                10                15

Leu Tyr Val Phe Leu Asn Leu Ser Ser Arg Lys Ser Ala Arg Leu Pro
      20                               25                30

Pro Gly Pro Thr Pro Trp Pro Ile Val Gly Asn Leu Pro His Leu Gly
      35                               40                45

```

- 161 -

Pro Ile Pro His His Ala Leu Ala Ala Leu Ala Lys Lys Tyr Gly Pro

50

55

60

Leu Met His Leu Arg Leu Gly Cys Val Asp Val Val Val Ala Ala Ser

65

70

75

80

Ala Ser Val Ala Ala Gln Phe Leu Lys Val His Asp Ala Asn Phe Ala

85

90

95

Ser Arg Pro Pro Asn Ser Gly Ala Lys His Val Ala Tyr Asn Tyr Gln

100

105

110

Asp Leu Val Phe Ala Pro Tyr Gly Pro Arg Trp Arg Leu Leu Arg Lys

115

120

125

Ile Cys Ser Val His Leu Phe Ser Ala Lys Ala Leu Asp Asp Phe Arg

130

135

140

His Val Arg Gln Glu Glu Val Ala Val Leu Thr Arg Val Leu Leu Ser

145

150

155

160

Ala Gly Asn Ser Pro Val Gln Leu Gly Gln Leu Leu Asn Val Cys Ala

165

170

175

Thr Asn Ala Leu Ala Arg Val Met Leu Gly Arg Arg Val Phe Gly Asp

180

185

190

Gly Ile Asp Arg Ser Ala Asn Glu Phe Lys Asp Met Val Val Glu Leu

195

200

205

Met Val Leu Ala Gly Glu Phe Asn Leu Gly Asp Phe Ile Pro Val Leu

210

215

220

Asp Leu Phe Asp Leu Gln Gly Ile Thr Lys Lys Met Lys Lys Leu His

225

230

235

240

Val Arg Phe Asp Ser Phe Leu Ser Lys Ile Val Glu Glu His Lys Thr

- 162 -

245	250	255
Ala Pro Gly Gly Leu Gly His Thr Asp Leu Leu Ser Thr Leu Ile Ser		
260	265	270
Leu Lys Asp Asp Ala Asp Ile Glu Gly Gly Lys Leu Thr Asp Thr Glu		
275	280	285
Ile Lys Ala Leu Leu Leu Asn Leu Phe Ala Ala Gly Thr Asp Thr Ser		
290	295	300
Ser Ser Thr Val Glu Trp Ala Ile Ala Glu Leu Ile Arg His Pro Gln		
305	310	315
Ile Leu Lys Gln Ala Arg Glu Glu Ile Asp Ala Val Val Gly Gln Asp		
325	330	335
Arg Leu Val Thr Glu Leu Asp Leu Ser Gln Leu Thr Tyr Leu Gln Ala		
340	345	350
Leu Val Lys Glu Val Phe Arg Leu His Pro Ser Thr Pro Leu Ser Leu		
355	360	365
Pro Arg Ile Ser Ser Glu Ser Cys Glu Val Asp Gly Tyr Tyr Ile Pro		
370	375	380
Lys Gly Ser Thr Leu Leu Val Asn Val Trp Ala Ile Ala Arg Asp Pro		
385	390	395
Lys Met Trp Ala Asp Pro Leu Glu Phe Arg Pro Ser Arg Phe Leu Pro		
405	410	415
Gly Gly Glu Lys Pro Gly Ala Asp Val Arg Gly Asn Asp Phe Glu Val		
420	425	430
Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala Gly Met Ser Leu Gly		
435	440	445

- 163 -

Leu Arg Met Val Gln Leu Leu Ile Ala Thr Leu Val Gln Thr Phe Asp

450

455

460

Trp Glu Leu Ala Asn Gly Leu Glu Pro Glu Met Leu Asn Met Glu Glu

465

470

475

480

Ala Tyr Gly Leu Thr Leu Gln Arg Ala Ala Pro Leu Met Val His Pro

485

490

495

Lys Pro Arg Leu Ala Pro His Val Tyr Glu Ser Ile

500

505

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1815 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 107..1631

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CTAAATTAAT TAATAAATAC ACACACGACG AGATGTATGT AATGTAATGT AATATTATTA 60

CATACATCAT CACCGAATAC GCACGCTACT ACCACTGCGA TTAGCC ATG AGT CCC 115
Met Ser Pro

1

TTA GCC TTG ATG ATC ATA AGT ACC TTA TTA GGG TTT CTC CTA TAC CAC 163

- 164 -

Leu Ala Leu Met Ile Ile Ser Thr Leu Leu Gly Phe Leu Leu Tyr His															
5				10				15							
TCT	CTT	CGC	TTA	CTA	CTC	TTC	TCC	GGC	CAA	GGT	CGC	CGA	CTA	CTA	CCA
20				25				30				35			
Ser	Leu	Arg	Leu	Leu	Leu	Phe	Ser	Gly	Gln	Gly	Arg	Arg	Leu	Leu	Pro
211															
CCA	GGT	CCA	CGC	CCG	TGG	CCG	CTG	GTG	GGA	AAT	CTC	CCG	CAC	TTA	GGC
40				45				50							
Pro	Gly	Pro	Arg	Pro	Trp	Pro	Leu	Val	Gly	Asn	Leu	Pro	His	Leu	Gly
259															
CCG	AAG	CCA	CAC	GCC	TCC	ATG	GCC	GAG	CTC	GCG	CGA	GCC	TAC	GGA	CCC
55				60				65							
Pro	Lys	Pro	His	Ala	Ser	Met	Ala	Glu	Leu	Ala	Arg	Ala	Tyr	Gly	Pro
307															
CTC	ATG	CAC	CTA	AAG	ATG	GGG	TTC	GTC	CAC	GTC	GTG	GTG	GCT	TCG	TCG
70				75				80							
Leu	Met	His	Leu	Lys	Met	Gly	Phe	Val	His	Val	Val	Val	Ala	Ser	Ser
355															
GCG	AGC	GCG	GCG	GAG	CAG	TGC	CTG	AGG	GTT	CAC	GAC	GCG	AAT	TTC	TTG
85				90				95							
Ala	Ser	Ala	Ala	Glu	Gln	Cys	Leu	Arg	Val	His	Asp	Ala	Asn	Phe	Leu
403															
AGC	AGG	CCA	CCC	AAC	TCC	GGC	GCC	AAG	CAC	GTC	GCT	TAC	AAC	TAC	GAG
100				105				110				115			
Ser	Arg	Pro	Pro	Asn	Ser	Gly	Ala	Lys	His	Val	Ala	Tyr	Asn	Tyr	Glu
451															
GAC	TTG	GTT	TTC	AGA	CCG	TAC	GGT	CCC	AAG	TGG	AGG	CTG	TTG	AGG	AAG
120				125				130							
Asp	Leu	Val	Phe	Arg	Pro	Tyr	Gly	Pro	Lys	Trp	Arg	Leu	Leu	Arg	Lys
499															
ATA	TGC	GCT	CAG	CAT	ATT	TTC	TCC	GTC	AAG	GCT	ATG	GAT	GAC	TTC	AGG
135				140				145							
Ile	Cys	Ala	Gln	His	Ile	Phe	Ser	Val	Lys	Ala	Met	Asp	Asp	Phe	Arg
547															
CGC	GTC	AGA	GAG	GAA	GAG	GTG	GCC	ATC	CTG	AGT	CGC	GCT	CTA	GCA	GGC
595															
Arg	Val	Arg	Glu	Glu	Glu	Val	Ala	Ile	Leu	Ser	Arg	Ala	Leu	Ala	Gly

- 165 -

150	155	160	
AAA AGG GCC GTA CCC ATA GGC CAA ATG CTC AAC GTG TGC GCC ACA AAC			643
Lys Arg Ala Val Pro Ile Gly Gln Met Leu Asn Val Cys Ala Thr Asn			
165	170	175	
GCC CTA TCT CGC GTC ATG ATG GGG CGG CGC GTG GTG GGC CAC GCG GAT			691
Ala Leu Ser Arg Val Met Met Gly Arg Arg Val Val Gly His Ala Asp			
180	185	190	195
GGA ACC AAC GAC GCC AAG GCG GAG GAG TTC AAA GCC ATG GTC GTC GAG			739
Gly Thr Asn Asp Ala Lys Ala Glu Glu Phe Lys Ala Met Val Val Glu			
200	205	210	
CTC ATG GTC CTC TCC GGC GTC TTC AAC ATC GGT GAT TTC ATC CCC TTC			787
Leu Met Val Leu Ser Gly Val Phe Asn Ile Gly Asp Phe Ile Pro Phe			
215	220	225	
CTC GAG CCT CTC GAC TTG CAG GGA GTG GCT TCC AAG ATG AAG AAA CTC			835
Leu Glu Pro Leu Asp Leu Gln Gly Val Ala Ser Lys Met Lys Lys Leu			
230	235	240	
CAC GCG CGG TTC GAT GCA TTC TTG ACC GAG ATT GTA CGA GAG CGT TGT			883
His Ala Arg Phe Asp Ala Phe Leu Thr Glu Ile Val Arg Glu Arg Cys			
245	250	255	
CAT GGG CAG ATC AAC AAC AGT GGT GCT CAT CAG GAT GAT TTG CTT AGC			931
His Gly Gln Ile Asn Asn Ser Gly Ala His Gln Asp Asp Leu Leu Ser			
260	265	270	275
ACG TTG ATT TCG TTC AAA GGG CTT GAC GAT GGC GAT GGT TCC AGG CTC			979
Thr Leu Ile Ser Phe Lys Gly Leu Asp Asp Gly Asp Gly Ser Arg Leu			
280	285	290	
ACT GAC ACA GAA ATC AAG GCG CTG CTC TTG AAC CTT TTG GAC ACG ACG			1027
Thr Asp Thr Glu Ile Lys Ala Leu Leu Leu Asn Leu Leu Asp Thr Thr			
295	300	305	

- 166 -

TCG	AGC	ACG	GTG	GAA	TGG	GCC	GTA	GCC	GAA	CTC	CTA	CGC	CAC	CCT	AAG	1075
Ser	Ser	Thr	Val	Glu	Trp	Ala	Val	Ala	Glu	Leu	Leu	Arg	His	Pro	Lys	
		310					315					320				
ACA	TTA	GCC	CAA	GTC	CGG	CAA	GAG	CTC	GAC	TCG	GTC	GTG	GGT	AAG	AAC	1123
Thr	Leu	Ala	Gln	Val	Arg	Gln	Glu	Leu	Asp	Ser	Val	Val	Gly	Lys	Asn	
		325				330						335				
AGG	CTC	GTG	TCC	GAG	ACC	GAT	CTG	AAT	CAG	CTG	CCC	TAT	CTA	CAA	GCT	1171
Arg	Leu	Val	Ser	Glu	Thr	Asp	Leu	Asn	Gln	Leu	Pro	Tyr	Leu	Gln	Ala	
340					345					350				355		
GTC	GTC	AAA	GAA	ACT	TTC	CGC	CTC	CAT	CCT	CCG	ACG	CCG	CTC	TCT	CTA	1219
Val	Val	Lys	Glu	Thr	Phe	Arg	Leu	His	Pro	Pro	Thr	Pro	Leu	Ser	Leu	
				360				365						370		
CCG	AGA	CTC	GCG	GAA	GAT	GAT	TGC	GAG	ATC	GAC	GGA	TAC	CTC	ATC	CCC	1267
Pro	Arg	Leu	Ala	Glu	Asp	Asp	Cys	Glu	Ile	Asp	Gly	Tyr	Leu	Ile	Pro	
		375					380					385				
AAG	GGC	TCG	ACC	CTT	CTG	GTG	AAC	GTT	TGG	GCC	ATA	GCC	CGC	GAT	CCC	1315
Lys	Gly	Ser	Thr	Leu	Leu	Val	Asn	Val	Trp	Ala	Ile	Ala	Arg	Asp	Pro	
		390					395					400				
AAG	GTT	TGG	GCC	GAT	CCG	TTG	GAG	TTT	AGG	CCC	GAA	CGA	TTC	TTG	ACG	1363
Lys	Val	Trp	Ala	Asp	Pro	Leu	Glu	Phe	Arg	Pro	Glu	Arg	Phe	Leu	Thr	
		405				410					415					
GGC	GGA	GAA	AAG	GCC	GAC	GTC	GAT	GTC	AAG	GGG	AAC	GAT	TTC	GAA	GTG	1411
Gly	Gly	Glu	Lys	Ala	Asp	Val	Asp	Val	Lys	Gly	Asn	Asp	Phe	Glu	Val	
420					425					430				435		
ATA	CCG	TTC	GGG	GCG	GGT	CGT	AGG	ATC	TGC	GCT	GGC	GTT	GGC	TTG	GGA	1459
Ile	Pro	Phe	Gly	Ala	Gly	Arg	Arg	Ile	Cys	Ala	Gly	Val	Gly	Leu	Gly	
				440					445					450		
ATA	CGT	ATG	GTC	CAA	CTG	TTG	ACG	GCG	AGT	TTG	ATC	CAT	GCA	TTC	GAT	1507

- 167 -

Ile Arg Met Val Gln Leu Leu Thr Ala Ser Leu Ile His Ala Phe Asp

455

460

465

CTG GAC CTT GCT AAT GGG CTT TTG GCC CAA AAT CTG AAC ATG GAA GAA 1555

Leu Asp Leu Ala Asn Gly Leu Leu Ala Gln Asn Leu Asn Met Glu Glu

470

475

480

GCA TAT GGG CTT ACG CTA CAA CGG GCT GAG CCT TTG TTG GTC CAC CCT 1603

Ala Tyr Gly Leu Thr Leu Gln Arg Ala Glu Pro Leu Leu Val His Pro

485

490

495

AGG CCG CGG TTG GCC ACT CAT GTC TAT T AATTAAATTA GGCCTAACT 1651

Arg Pro Arg Leu Ala Thr His Val Tyr

500

505

ACGATGAATG ACCCATTTAA CGTTAATAAG AGTTTTCAAT TTATGTGAGT TTGCATGGTA 1711

TGGTATGGTA TGGTGCTTGT AATAAATTGT ATCTGTTAGG TGTGTTTCATT GATGATAAAT 1771

CTAGTTTGTA CTGCTGCTCA AAAAAAAAAA AAAAAAAAAA AAAA 1815

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 508 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met Ser Pro Leu Ala Leu Met Ile Ile Ser Thr Leu Leu Gly Phe Leu

1

5

10

15

Leu Tyr His Ser Leu Arg Leu Leu Leu Phe Ser Gly Gln Gly Arg Arg

- 168 -

20	25	30
Leu Leu Pro Pro Gly Pro Arg Pro Trp Pro Leu Val Gly Asn Leu Pro		
35	40	45
His Leu Gly Pro Lys Pro His Ala Ser Met Ala Glu Leu Ala Arg Ala		
50	55	60
Tyr Gly Pro Leu Met His Leu Lys Met Gly Phe Val His Val Val Val		
65	70	75 80
Ala Ser Ser Ala Ser Ala Ala Glu Gln Cys Leu Arg Val His Asp Ala		
85	90	95
Asn Phe Leu Ser Arg Pro Pro Asn Ser Gly Ala Lys His Val Ala Tyr		
100	105	110
Asn Tyr Glu Asp Leu Val Phe Arg Pro Tyr Gly Pro Lys Trp Arg Leu		
115	120	125
Leu Arg Lys Ile Cys Ala Gln His Ile Phe Ser Val Lys Ala Met Asp		
130	135	140
Asp Phe Arg Arg Val Arg Glu Glu Glu Val Ala Ile Leu Ser Arg Ala		
145	150	155 160
Leu Ala Gly Lys Arg Ala Val Pro Ile Gly Gln Met Leu Asn Val Cys		
165	170	175
Ala Thr Asn Ala Leu Ser Arg Val Met Met Gly Arg Arg Val Val Gly		
180	185	190
His Ala Asp Gly Thr Asn Asp Ala Lys Ala Glu Glu Phe Lys Ala Met		
195	200	205
Val Val Glu Leu Met Val Leu Ser Gly Val Phe Asn Ile Gly Asp Phe		
210	215	220

- 169 -

Ile	Pro	Phe	Leu	Glu	Pro	Leu	Asp	Leu	Gln	Gly	Val	Ala	Ser	Lys	Met
225					230					235					240
Lys	Lys	Leu	His	Ala	Arg	Phe	Asp	Ala	Phe	Leu	Thr	Glu	Ile	Val	Arg
				245						250				255	
Glu	Arg	Cys	His	Gly	Gln	Ile	Asn	Asn	Ser	Gly	Ala	His	Gln	Asp	Asp
			260						265					270	
Leu	Leu	Ser	Thr	Leu	Ile	Ser	Phe	Lys	Gly	Leu	Asp	Asp	Gly	Asp	Gly
		275						280					285		
Ser	Arg	Leu	Thr	Asp	Thr	Glu	Ile	Lys	Ala	Leu	Leu	Leu	Asn	Leu	Leu
	290						295					300			
Asp	Thr	Thr	Ser	Ser	Thr	Val	Glu	Trp	Ala	Val	Ala	Glu	Leu	Leu	Arg
305					310					315					320
His	Pro	Lys	Thr	Leu	Ala	Gln	Val	Arg	Gln	Glu	Leu	Asp	Ser	Val	Val
				325						330				335	
Gly	Lys	Asn	Arg	Leu	Val	Ser	Glu	Thr	Asp	Leu	Asn	Gln	Leu	Pro	Tyr
			340						345				350		
Leu	Gln	Ala	Val	Val	Lys	Glu	Thr	Phe	Arg	Leu	His	Pro	Pro	Thr	Pro
		355							360				365		
Leu	Ser	Leu	Pro	Arg	Leu	Ala	Glu	Asp	Asp	Cys	Glu	Ile	Asp	Gly	Tyr
	370							375				380			
Leu	Ile	Pro	Lys	Gly	Ser	Thr	Leu	Leu	Val	Asn	Val	Trp	Ala	Ile	Ala
385						390					395				400
Arg	Asp	Pro	Lys	Val	Trp	Ala	Asp	Pro	Leu	Glu	Phe	Arg	Pro	Glu	Arg
				405						410				415	
Phe	Leu	Thr	Gly	Gly	Glu	Lys	Ala	Asp	Val	Asp	Val	Lys	Gly	Asn	Asp

- 170 -

420	425	430
Phe Glu Val Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala Gly Val		
435	440	445
Gly Leu Gly Ile Arg Met Val Gln Leu Leu Thr Ala Ser Leu Ile His		
450	455	460
Ala Phe Asp Leu Asp Leu Ala Asn Gly Leu Leu Ala Gln Asn Leu Asn		
465	470	475 480
Met Glu Glu Ala Tyr Gly Leu Thr Leu Gln Arg Ala Glu Pro Leu Leu		
485	490	495
Val His Pro Arg Pro Arg Leu Ala Thr His Val Tyr		
500	505	

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1824 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..1553

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

G AGC TTA ACC TTA ATT TTC TGC ACT TTA GTT TTT GCA ATC TTT CTA
 Ser Leu Thr Leu Ile Phe Cys Thr Leu Val Phe Ala Ile Phe Leu

- 171 -

1	5	10	15	
TAT TTT CTT ATT CTC AGG GTG AAA CAG CGT TAC CCT TTA CCT CTC CCA				94
Tyr Phe Leu Ile Leu Arg Val Lys Gln Arg Tyr Pro Leu Pro Leu Pro				
	20	25	30	
CCC GGA CCA AAA CCA TGG CCG GTG TTA GGA AAC CTT CCC CAC CTG GGC				142
Pro Gly Pro Lys Pro Trp Pro Val Leu Gly Asn Leu Pro His Leu Gly				
	35	40	45	
AAG AAG CCT CAC CAG TCG ATT GCG GCC ATG GCT GAG AGG TAC GGC CCC				190
Lys Lys Pro His Gln Ser Ile Ala Ala Met Ala Glu Arg Tyr Gly Pro				
	50	55	60	
CTC ATG CAC CTC CGC CTA GGA TTC GTG GAC GTG GTT GTG GCC GCC TCC				238
Leu Met His Leu Arg Leu Gly Phe Val Asp Val Val Val Ala Ala Ser				
	65	70	75	
GCC GCC GTG GCC GCT CAG TTC TTG AAA GTT CAC GAC TCG AAC TTC TCC				286
Ala Ala Val Ala Ala Gln Phe Leu Lys Val His Asp Ser Asn Phe Ser				
	80	85	90	95
AAC CGG CCG CCG AAC TCC GGC GCG GAA CAC ATT GCT TAT AAC TAT CAA				334
Asn Arg Pro Pro Asn Ser Gly Ala Glu His Ile Ala Tyr Asn Tyr Gln				
	100	105	110	
GAC CTC GTC TTC GCG CCC TAC GGC CCG CGG TGG CGC ATG CTT AGG AAG				382
Asp Leu Val Phe Ala Pro Tyr Gly Pro Arg Trp Arg Met Leu Arg Lys				
	115	120	125	
ATC ACC TCC GTG CAT CTC TTC TCG GCC AAG GCG TTG GAT GAC TTC TGC				430
Ile Thr Ser Val His Leu Phe Ser Ala Lys Ala Leu Asp Asp Phe Cys				
	130	135	140	
CAT GTT CGC CAG GAA GAG GTT GCA ACT CTG ACA CGC AGT CTA GCA AGT				478
His Val Arg Gln Glu Glu Val Ala Thr Leu Thr Arg Ser Leu Ala Ser				
	145	150	155	

- 172 -

GCA GGC AAA ACT CCA GTA AAA CTA GGG CAG TTA CTA AAC GTG TGC ACC	526
Ala Gly Lys Thr Pro Val Lys Leu Gly Gln Leu Leu Asn Val Cys Thr	
160 165 170 175	
ACG AAC GCC CTA GCT CGT GTA ATG CTA GGG CGG AAG GTC TTT AAT GAC	574
Thr Asn Ala Leu Ala Arg Val Met Leu Gly Arg Lys Val Phe Asn Asp	
180 185 190	
GGA GGT AGC AAG AGC GAC CCA AAG GCG GAG GAG TTC AAG TCG ATG GTG	622
Gly Gly Ser Lys Ser Asp Pro Lys Ala Glu Glu Phe Lys Ser Met Val	
195 200 205	
GAG GAG ATG ATG GTG TTG GCC GGA AGT TTT AAC ATC GGC GAT TTC ATT	670
Glu Glu Met Met Val Leu Ala Gly Ser Phe Asn Ile Gly Asp Phe Ile	
210 215 220	
CCG GTC TTG GGT TGG TTT GAC GTT CAG GGT ATC GTA GGG AAG ATG AAG	718
Pro Val Leu Gly Trp Phe Asp Val Gln Gly Ile Val Gly Lys Met Lys	
225 230 235	
AAA CTA CAC GCG CGT TTT GAT GCG TTC TTG AAC ACC ATT CTA GAG GAA	766
Lys Leu His Ala Arg Phe Asp Ala Phe Leu Asn Thr Ile Leu Glu Glu	
240 245 250 255	
CAC AAA TGT GTC AAC AAT CAA CAC ACG ACG TTG TCG AAA GAT GTG GAC	814
His Lys Cys Val Asn Asn Gln His Thr Thr Leu Ser Lys Asp Val Asp	
260 265 270	
TTC TTG AGC ACC CTA ATT AGG CTC AAA GAT AAT GGG GCT GAT ATG GAT	862
Phe Leu Ser Thr Leu Ile Arg Leu Lys Asp Asn Gly Ala Asp Met Asp	
275 280 285	
TGT GAA GAG GGA AAA CTC ACC GAC ACT GAA ATT AAG GCT TTG CTC TTG	910
Cys Glu Glu Gly Lys Leu Thr Asp Thr Glu Ile Lys Ala Leu Leu Leu	
290 295 300	
AAC CTG TTC ACA GCT GGG ACT GAT ACA TCA TCT AGC ACT GTG GAG TGG	958

- 174 -

450	455	460	
TTG GTT GCA ACT TTG GTG CAT GCT TTT GAT TGG GAT TTG GTG AAT GGA			1438
Leu Val Ala Thr Leu Val His Ala Phe Asp Trp Asp Leu Val Asn Gly			
465	470	475	
CAA TCT GTA GAG ACG CTC AAT ATG GAG GAA GCT TAT GGT CTC ACC CTT			1486
Gln Ser Val Glu Thr Leu Asn Met Glu Glu Ala Tyr Gly Leu Thr Leu			
480	485	490	495
CAA CGA GCT GTT CCT TTG ATG TTG CAT CCA AAG CCC AGA TTA CAA CCA			1534
Gln Arg Ala Val Pro Leu Met Leu His Pro Lys Pro Arg Leu Gln Pro			
500	505	510	
CAT CTC TAT ACT CTC AAT T AAATTGCAAT TTGATTTTGG TGATTATACA			1583
His Leu Tyr Thr Leu Asn			
515			
ATTATAATCG AGGGACATAG GATCCCCATT TATTTATATT CAGTTATAAG AGACTTCCAA			1643
CAAAGGTCTA GCTTTCGACC TTAAAAGTTG TAAAAGAGGT CCTACATATG TAAAAGCCCG			1703
CCAAAGGAAA ACTGGTTGTA TTCAATTCCG CTAGGCCTTG TCCGAAAGAC CTCATGAAGA			1763
CTACAAAGGT CATATATAAT GGTAACCCA GTGTATTTGT TGTAAAAAAA AAAAAAAAAA			1823
A			1824

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 517 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- 175 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Ser	Leu	Thr	Leu	Ile	Phe	Cys	Thr	Leu	Val	Phe	Ala	Ile	Phe	Leu	Tyr
1				5					10					15	
Phe	Leu	Ile	Leu	Arg	Val	Lys	Gln	Arg	Tyr	Pro	Leu	Pro	Leu	Pro	Pro
			20					25					30		
Gly	Pro	Lys	Pro	Trp	Pro	Val	Leu	Gly	Asn	Leu	Pro	His	Leu	Gly	Lys
		35					40						45		
Lys	Pro	His	Gln	Ser	Ile	Ala	Ala	Met	Ala	Glu	Arg	Tyr	Gly	Pro	Leu
	50					55						60			
Met	His	Leu	Arg	Leu	Gly	Phe	Val	Asp	Val	Val	Val	Ala	Ala	Ser	Ala
65					70					75					80
Ala	Val	Ala	Ala	Gln	Phe	Leu	Lys	Val	His	Asp	Ser	Asn	Phe	Ser	Asn
				85						90				95	
Arg	Pro	Pro	Asn	Ser	Gly	Ala	Glu	His	Ile	Ala	Tyr	Asn	Tyr	Gln	Asp
			100						105					110	
Leu	Val	Phe	Ala	Pro	Tyr	Gly	Pro	Arg	Trp	Arg	Met	Leu	Arg	Lys	Ile
			115						120				125		
Thr	Ser	Val	His	Leu	Phe	Ser	Ala	Lys	Ala	Leu	Asp	Asp	Phe	Cys	His
	130						135					140			
Val	Arg	Gln	Glu	Glu	Val	Ala	Thr	Leu	Thr	Arg	Ser	Leu	Ala	Ser	Ala
145					150					155				160	
Gly	Lys	Thr	Pro	Val	Lys	Leu	Gly	Gln	Leu	Leu	Asn	Val	Cys	Thr	Thr
				165					170					175	
Asn	Ala	Leu	Ala	Arg	Val	Met	Leu	Gly	Arg	Lys	Val	Phe	Asn	Asp	Gly
		180						185						190	

- 176 -

Gly Ser Lys Ser Asp Pro Lys Ala Glu Glu Phe Lys Ser Met Val Glu

195

200

205

Glu Met Met Val Leu Ala Gly Ser Phe Asn Ile Gly Asp Phe Ile Pro

210

215

220

Val Leu Gly Trp Phe Asp Val Gln Gly Ile Val Gly Lys Met Lys Lys

225

230

235

240

Leu His Ala Arg Phe Asp Ala Phe Leu Asn Thr Ile Leu Glu Glu His

245

250

255

Lys Cys Val Asn Asn Gln His Thr Thr Leu Ser Lys Asp Val Asp Phe

260

265

270

Leu Ser Thr Leu Ile Arg Leu Lys Asp Asn Gly Ala Asp Met Asp Cys

275

280

285

Glu Glu Gly Lys Leu Thr Asp Thr Glu Ile Lys Ala Leu Leu Leu Asn

290

295

300

Leu Phe Thr Ala Gly Thr Asp Thr Ser Ser Ser Thr Val Glu Trp Ala

305

310

315

320

Ile Ala Glu Leu Leu Arg Asn Pro Lys Ile Leu Asn Gln Ala Gln Gln

325

330

335

Glu Leu Asp Leu Val Val Gly Gln Asn Gln Leu Val Thr Glu Ser Asp

340

345

350

Leu Thr Asp Leu Pro Phe Leu Gln Ala Ile Val Lys Glu Thr Phe Arg

355

360

365

Leu His Pro Ser Thr Pro Leu Ser Leu Pro Arg Met Gly Ala Gln Gly

370

375

380

Cys Glu Ile Asn Gly Tyr Phe Ile Pro Lys Gly Ala Thr Leu Leu Val

- 177 -

385	390	395	400
Asn Val Trp Ala Ile Ala Arg Asp Pro Asn Val Trp Thr Asn Pro Leu	405	410	415
Glu Phe Asn Pro His Arg Phe Leu Pro Gly Gly Glu Lys Pro Asn Val	420	425	430
Asp Ile Lys Gly Asn Asp Phe Glu Val Ile Pro Phe Gly Ala Gly Arg	435	440	445
Arg Ile Cys Ser Gly Met Ser Leu Gly Ile Arg Met Val His Leu Leu	450	455	460
Val Ala Thr Leu Val His Ala Phe Asp Trp Asp Leu Val Asn Gly Gln	465	470	475
Ser Val Glu Thr Leu Asn Met Glu Glu Ala Tyr Gly Leu Thr Leu Gln	485	490	495
Arg Ala Val Pro Leu Met Leu His Pro Lys Pro Arg Leu Gln Pro His	500	505	510
Leu Tyr Thr Leu Asn	515		

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1667 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- 178 -

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..1429

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CCC	ATC	CTC	GGA	AAC	ATC	CCC	CAT	CTC	GGC	TCC	AAA	CCG	CAC	CAA	ACA	48
Pro	Ile	Leu	Gly	Asn	Ile	Pro	His	Leu	Gly	Ser	Lys	Pro	His	Gln	Thr	
1				5					10					15		
CTC	GCG	GAA	ATG	GCG	AAA	ACC	TAC	GGT	CCG	CTC	ATG	CAC	TTG	AAG	TTC	96
Leu	Ala	Glu	Met	Ala	Lys	Thr	Tyr	Gly	Pro	Leu	Met	His	Leu	Lys	Phe	
				20				25					30			
GGG	CTT	AAG	GAC	GCG	GTG	GTG	GCG	TCG	TCT	GCG	TCG	GTG	GCA	GAG	CAG	144
Gly	Leu	Lys	Asp	Ala	Val	Val	Ala	Ser	Ser	Ala	Ser	Val	Ala	Glu	Gln	
				35				40					45			
TTT	CTG	AAG	AAA	CAC	GAC	GTG	AAT	TTC	TCG	AAC	CGG	CCG	CCA	AAC	TCC	192
Phe	Leu	Lys	Lys	His	Asp	Val	Asn	Phe	Ser	Asn	Arg	Pro	Pro	Asn	Ser	
				50				55				60				
GGG	GCC	AAA	CAT	ATA	GCT	TAT	AAC	TAT	CAG	GAC	CTG	GTA	TTC	GCT	CCC	240
Gly	Ala	Lys	His	Ile	Ala	Tyr	Asn	Tyr	Gln	Asp	Leu	Val	Phe	Ala	Pro	
65					70					75					80	
TAT	GGA	CCC	CGG	TGG	CGG	TTG	CTT	AGG	AAA	ATC	TGT	TCC	GTC	CAT	CTT	288
Tyr	Gly	Pro	Arg	Trp	Arg	Leu	Leu	Arg	Lys	Ile	Cys	Ser	Val	His	Leu	
					85					90				95		
TTC	TCG	TCT	AAG	GCC	TTG	GAT	GAC	TTT	CAG	CAT	GTT	CGA	CAT	GAG	GAG	336
Phe	Ser	Ser	Lys	Ala	Leu	Asp	Asp	Phe	Gln	His	Val	Arg	His	Glu	Glu	
				100				105					110			
ATA	TGC	ATC	CTT	ATA	CGA	GCA	ATA	GCG	AGT	GGC	GGT	CAT	GCT	CCG	GTG	384
Ile	Cys	Ile	Leu	Ile	Arg	Ala	Ile	Ala	Ser	Gly	Gly	His	Ala	Pro	Val	

- 179 -

115	120	125	
AAT TTA GGC AAG TTA TTA GGA GTG TGC ACA ACC AAT GCC CTG GCA AGA			432
Asn Leu Gly Lys Leu Leu Gly Val Cys Thr Thr Asn Ala Leu Ala Arg			
130	135	140	
GTG ATG CTT GGA AGA AGA GTA TTC GAA GGC GAC GGC GGC GAG AAT CCG			480
Val Met Leu Gly Arg Arg Val Phe Glu Gly Asp Gly Gly Glu Asn Pro			
145	150	155	160
CAT GCC GAC GAG TTT AAA TCA ATG GTG GTG GAG ATT ATG GTG TTA GCC			528
His Ala Asp Glu Phe Lys Ser Met Val Val Glu Ile Met Val Leu Ala			
165	170	175	
GGT GCA TTC AAC TTG GGT GAT TTC ATC CCG GTT CTA GAT TGG TTC GAT			576
Gly Ala Phe Asn Leu Gly Asp Phe Ile Pro Val Leu Asp Trp Phe Asp			
180	185	190	
TTG CAA GGA ATT GCT GGT AAA ATG AAG AAA CTT CAT GCC CGT TTC GAC			624
Leu Gln Gly Ile Ala Gly Lys Met Lys Lys Leu His Ala Arg Phe Asp			
195	200	205	
AAG TTT TTA AAT GGG ATC CTA GAA GAT CGT AAA TCT AAC GGC TCT AAT			672
Lys Phe Leu Asn Gly Ile Leu Glu Asp Arg Lys Ser Asn Gly Ser Asn			
210	215	220	
GGA GCT GAA CAA TAC GTG GAC TTG CTC AGT GTG TTG ATC TCT CTT CAA			720
Gly Ala Glu Gln Tyr Val Asp Leu Leu Ser Val Leu Ile Ser Leu Gln			
225	230	235	240
GAT AGT AAT ATC GAC GGT GGT GAC GAA GGA ACC AAA CTC ACA GAT ACT			768
Asp Ser Asn Ile Asp Gly Gly Asp Glu Gly Thr Lys Leu Thr Asp Thr			
245	250	255	
GAA ATC AAA GCT CTC CTT TTG AAC TTG TTC ATA GCC GGA ACA GAC ACT			816
Glu Ile Lys Ala Leu Leu Leu Asn Leu Phe Ile Ala Gly Thr Asp Thr			
260	265	270	

- 180 -

TCA TCA AGT ACT GTA GAA TGG GCC ATG GCA GAA CTA ATC CGA AAC CCA	864
Ser Ser Ser Thr Val Glu Trp Ala Met Ala Glu Leu Ile Arg Asn Pro	
275 280 285	
AAG TTA CTA GTC CAA GCC CAA GAA GAG CTA GAC AGA GTA GTC GGG CCG	912
Lys Leu Leu Val Gln Ala Gln Glu Glu Leu Asp Arg Val Val Gly Pro	
290 295 300	
AAC CGA TTC GTA ACC GAA TCT GAT CTT CCT CAA CTG ACA TTC CTT CAA	960
Asn Arg Phe Val Thr Glu Ser Asp Leu Pro Gln Leu Thr Phe Leu Gln	
305 310 315 320	
GCC GTC ATC AAA GAG ACT TTC AGG CTT CAT CCA TCC ACC CCA CTC TCT	1008
Ala Val Ile Lys Glu Thr Phe Arg Leu His Pro Ser Thr Pro Leu Ser	
325 330 335	
CTT CCA CGA ATG GCG GCG GAG GAC TGT GAG ATC AAT GGG TAT TAT GTC	1056
Leu Pro Arg Met Ala Ala Glu Asp Cys Glu Ile Asn Gly Tyr Tyr Val	
340 345 350	
TCA GAA GGT TCG ACA TTG CTC GTC AAT GTG TGG GCC ATA GCT CGT GAT	1104
Ser Glu Gly Ser Thr Leu Leu Val Asn Val Trp Ala Ile Ala Arg Asp	
355 360 365	
CCA AAT GCG TGG GCC AAT CCA CTA GAT TTC AAC CCG ACT CGT TTC TTG	1152
Pro Asn Ala Trp Ala Asn Pro Leu Asp Phe Asn Pro Thr Arg Phe Leu	
370 375 380	
GCC GGT GGA GAG AAG CCT AAT GTT GAT GTT AAA GGG AAT GAT TTT GAA	1200
Ala Gly Gly Glu Lys Pro Asn Val Asp Val Lys Gly Asn Asp Phe Glu	
385 390 395 400	
GTG ATA CCT TTC GGT GCT GGG CGC AGG ATA TGT GCC GGA ATG AGC TTA	1248
Val Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala Gly Met Ser Leu	
405 410 415	
GGT ATA CGG ATG GTT CAA CTA GTA ACG GCT TCG TTA GTT CAT TCG TTT	1296

- 181 -

Gly Ile Arg Met Val Gln Leu Val Thr Ala Ser Leu Val His Ser Phe

420

425

430

GAT TGG GCT TTG TTG GAT GGA CTT AAA CCC GAG AAG CTT GAC ATG GAG 1344

Asp Trp Ala Leu Leu Asp Gly Leu Lys Pro Glu Lys Leu Asp Met Glu

435

440

445

GAA GGT TAT GGA CTA ACG CTT CAA CGA GCT TCA CCT TTA ATC GTC CAT 1392

Glu Gly Tyr Gly Leu Thr Leu Gln Arg Ala Ser Pro Leu Ile Val His

450

455

460

CCA AAG CCG AGG CTC TCG GCT CAA GTT TAT TGT ATG T AACAAGTTTG 1439

Pro Lys Pro Arg Leu Ser Ala Gln Val Tyr Cys Met

465

470

475

TGAAGCCAGT CTGATTCAG TTGGATTTGT AGTTATTTTA TGATCATTG GTATTTTATT 1499

TTGTATTTTCG GTTGAATACA ATAAAGGGAA GGTGGATCGT CTGCTGTATA ATAGCGACGT 1559

TTTAACGTGT TGTGATAGTA CCGTGTTTTA CTAAAACGAT GTCGTTTGAT TTTTATAGT 1619

ATTAAAAAAA TAAACAGCTG GATTTTGAAC CAAAAAAA AAAAAAAA 1667

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 476 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Pro Ile Leu Gly Asn Ile Pro His Leu Gly Ser Lys Pro His Gln Thr

1

5

10

15

- 182 -

Leu	Ala	Glu	Met	Ala	Lys	Thr	Tyr	Gly	Pro	Leu	Met	His	Leu	Lys	Phe
			20					25					30		

Gly	Leu	Lys	Asp	Ala	Val	Val	Ala	Ser	Ser	Ala	Ser	Val	Ala	Glu	Gln
		35						40					45		

Phe	Leu	Lys	Lys	His	Asp	Val	Asn	Phe	Ser	Asn	Arg	Pro	Pro	Asn	Ser
		50					55					60			

Gly	Ala	Lys	His	Ile	Ala	Tyr	Asn	Tyr	Gln	Asp	Leu	Val	Phe	Ala	Pro
65					70					75					80

Tyr	Gly	Pro	Arg	Trp	Arg	Leu	Leu	Arg	Lys	Ile	Cys	Ser	Val	His	Leu
				85					90					95	

Phe	Ser	Ser	Lys	Ala	Leu	Asp	Asp	Phe	Gln	His	Val	Arg	His	Glu	Glu
			100					105					110		

Ile	Cys	Ile	Leu	Ile	Arg	Ala	Ile	Ala	Ser	Gly	Gly	His	Ala	Pro	Val
		115						120					125		

Asn	Leu	Gly	Lys	Leu	Leu	Gly	Val	Cys	Thr	Thr	Asn	Ala	Leu	Ala	Arg
		130					135					140			

Val	Met	Leu	Gly	Arg	Arg	Val	Phe	Glu	Gly	Asp	Gly	Gly	Glu	Asn	Pro
145						150				155					160

His	Ala	Asp	Glu	Phe	Lys	Ser	Met	Val	Val	Glu	Ile	Met	Val	Leu	Ala
			165							170				175	

Gly	Ala	Phe	Asn	Leu	Gly	Asp	Phe	Ile	Pro	Val	Leu	Asp	Trp	Phe	Asp
			180					185					190		

Leu	Gln	Gly	Ile	Ala	Gly	Lys	Met	Lys	Lys	Leu	His	Ala	Arg	Phe	Asp
		195						200					205		

Lys	Phe	Leu	Asn	Gly	Ile	Leu	Glu	Asp	Arg	Lys	Ser	Asn	Gly	Ser	Asn
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- 183 -

210	215	220	
Gly Ala Glu Gln Tyr Val Asp Leu Leu Ser Val Leu Ile Ser Leu Gln			
225	230	235	240
Asp Ser Asn Ile Asp Gly Gly Asp Glu Gly Thr Lys Leu Thr Asp Thr			
245	250	255	
Glu Ile Lys Ala Leu Leu Leu Asn Leu Phe Ile Ala Gly Thr Asp Thr			
260	265	270	
Ser Ser Ser Thr Val Glu Trp Ala Met Ala Glu Leu Ile Arg Asn Pro			
275	280	285	
Lys Leu Leu Val Gln Ala Gln Glu Glu Leu Asp Arg Val Val Gly Pro			
290	295	300	
Asn Arg Phe Val Thr Glu Ser Asp Leu Pro Gln Leu Thr Phe Leu Gln			
305	310	315	320
Ala Val Ile Lys Glu Thr Phe Arg Leu His Pro Ser Thr Pro Leu Ser			
325	330	335	
Leu Pro Arg Met Ala Ala Glu Asp Cys Glu Ile Asn Gly Tyr Tyr Val			
340	345	350	
Ser Glu Gly Ser Thr Leu Leu Val Asn Val Trp Ala Ile Ala Arg Asp			
355	360	365	
Pro Asn Ala Trp Ala Asn Pro Leu Asp Phe Asn Pro Thr Arg Phe Leu			
370	375	380	
Ala Gly Gly Glu Lys Pro Asn Val Asp Val Lys Gly Asn Asp Phe Glu			
385	390	395	400
Val Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala Gly Met Ser Leu			
405	410	415	

- 184 -

Gly Ile Arg Met Val Gln Leu Val Thr Ala Ser Leu Val His Ser Phe

420

425

430

Asp Trp Ala Leu Leu Asp Gly Leu Lys Pro Glu Lys Leu Asp Met Glu

435

440

445

Glu Gly Tyr Gly Leu Thr Leu Gln Arg Ala Ser Pro Leu Ile Val His

450

455

460

Pro Lys Pro Arg Leu Ser Ala Gln Val Tyr Cys Met

465

470

475

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1214 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 2..1091

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

T CGC ATC CTC ACG CGA TCT ATA GCG AGT GCT GGG GAA AAT CCG ATT 46

Arg Ile Leu Thr Arg Ser Ile Ala Ser Ala Gly Glu Asn Pro Ile

1

5

10

15

AAC TTA GGT CAA TTA CTC GGG GTG TGT ACC ACA AAT GCT CTG GCG AGA 94

Asn Leu Gly Gln Leu Leu Gly Val Cys Thr Thr Asn Ala Leu Ala Arg

20

25

30

- 185 -

GTG	ATG	CTT	GGA	AGG	AGG	GTA	TTC	GGC	GAT	GGG	AGC	GGC	GGC	GTA	GAT	142
Val	Met	Leu	Gly	Arg	Arg	Val	Phe	Gly	Asp	Gly	Ser	Gly	Gly	Val	Asp	
			35					40					45			
CCT	CAG	GCG	GAC	GAG	TTC	AAA	TCC	ATG	GTG	GTG	GAA	ATC	ATG	GTG	TTG	190
Pro	Gln	Ala	Asp	Glu	Phe	Lys	Ser	Met	Val	Val	Glu	Ile	Met	Val	Leu	
			50				55					60				
GCC	GGC	GCG	TTT	AAT	CTA	GGT	GAT	TTT	ATT	CCC	GCT	CTT	GAT	TGG	TTC	238
Ala	Gly	Ala	Phe	Asn	Leu	Gly	Asp	Phe	Ile	Pro	Ala	Leu	Asp	Trp	Phe	
			65				70					75				
GAT	CTG	CAG	GGA	ATT	ACG	GCA	AAA	ATG	AAG	AAA	GTT	CAC	GCT	CGT	TTC	286
Asp	Leu	Gln	Gly	Ile	Thr	Ala	Lys	Met	Lys	Lys	Val	His	Ala	Arg	Phe	
	80				85					90					95	
GAT	GCG	TTC	TTA	GAC	GCG	ATC	CTT	GAG	GAG	CAC	AAA	TCC	AAC	GGC	TCT	334
Asp	Ala	Phe	Leu	Asp	Ala	Ile	Leu	Glu	Glu	His	Lys	Ser	Asn	Gly	Ser	
					100					105				110		
CGC	GGA	GCT	AAG	CAA	CAC	GTT	GAC	TTG	CTG	AGT	ATG	TTG	ATC	TCC	CTT	382
Arg	Gly	Ala	Lys	Gln	His	Val	Asp	Leu	Leu	Ser	Met	Leu	Ile	Ser	Leu	
			115					120					125			
CAA	GAT	AAT	AAC	ATT	GAT	GGT	GAA	AGT	GGC	GCC	AAA	CTC	ACT	GAT	ACA	430
Gln	Asp	Asn	Asn	Ile	Asp	Gly	Glu	Ser	Gly	Ala	Lys	Leu	Thr	Asp	Thr	
			130				135					140				
GAA	ATC	AAA	GCT	TTG	CTT	CTG	AAC	TTG	TTC	ACG	GCT	GGA	ACA	GAC	ACG	478
Glu	Ile	Lys	Ala	Leu	Leu	Leu	Asn	Leu	Phe	Thr	Ala	Gly	Thr	Asp	Thr	
	145					150					155					
TCA	TCA	AGT	ACT	GTG	GAG	TGG	GCA	ATC	GCA	GAG	CTA	ATC	CGA	AAC	CCA	526
Ser	Ser	Ser	Thr	Val	Glu	Trp	Ala	Ile	Ala	Glu	Leu	Ile	Arg	Asn	Pro	
	160					165				170				175		
GAA	GTA	TTG	GTT	CAA	GCC	CAA	CAA	GAG	CTC	GAT	AGA	GTA	GTT	GGG	CCA	574

- 186 -

Glu	Val	Leu	Val	Gln	Ala	Gln	Gln	Glu	Leu	Asp	Arg	Val	Val	Gly	Pro		
				180				185				190					
AGT	CGT	CTT	GTG	ACC	GAA	TCT	GAT	CTG	CCT	CAA	TTG	GCA	TTC	CTT	CAA	622	
Ser	Arg	Leu	Val	Thr	Glu	Ser	Asp	Leu	Pro	Gln	Leu	Ala	Phe	Leu	Gln		
				195				200				205					
GCT	GTC	ATC	AAA	GAG	ACT	TTC	AGA	CTT	CAT	CCA	TCC	ACT	CCA	CTC	TCT	670	
Ala	Val	Ile	Lys	Glu	Thr	Phe	Arg	Leu	His	Pro	Ser	Thr	Pro	Leu	Ser		
				210				215				220					
CTT	CCA	CGA	ATG	GCT	TCA	GAG	GGT	TGT	GAA	ATC	AAT	GGA	TAC	TCC	ATC	718	
Leu	Pro	Arg	Met	Ala	Ser	Glu	Gly	Cys	Glu	Ile	Asn	Gly	Tyr	Ser	Ile		
				225				230				235					
CCA	AAG	GGT	TCG	ACA	TTG	CTC	GTT	AAC	GTA	TGG	TCC	ATA	GCC	CGT	GAT	766	
Pro	Lys	Gly	Ser	Thr	Leu	Leu	Val	Asn	Val	Trp	Ser	Ile	Ala	Arg	Asp		
				240				245				250				255	
CCT	AGT	ATA	TGG	GCC	GAC	CCA	TTA	GAA	TTT	AGG	CCG	GCA	CGT	TTC	TTG	814	
Pro	Ser	Ile	Trp	Ala	Asp	Pro	Leu	Glu	Phe	Arg	Pro	Ala	Arg	Phe	Leu		
				260				265				270					
CCC	GGC	GGA	GAA	AAG	CCC	AAT	GTT	GAT	GTG	AGA	GGC	AAT	GAT	TTT	GAG	862	
Pro	Gly	Gly	Glu	Lys	Pro	Asn	Val	Asp	Val	Arg	Gly	Asn	Asp	Phe	Glu		
				275				280				285					
GTC	ATA	CCA	TTT	GGT	GCT	GGA	CGT	AGG	ATA	TGT	GCT	GGA	ATG	AGC	TTG	910	
Val	Ile	Pro	Phe	Gly	Ala	Gly	Arg	Arg	Ile	Cys	Ala	Gly	Met	Ser	Leu		
				290				295				300					
GGT	TTA	AGA	ATG	GTT	CAA	CTT	TCG	ACA	GCT	ACT	TTG	GTT	CAT	TCG	TTT	958	
Gly	Leu	Arg	Met	Val	Gln	Leu	Ser	Thr	Ala	Thr	Leu	Val	His	Ser	Phe		
				305				310				315					
AAT	TGG	GAT	TTG	CTG	AAT	GGG	ATG	AGC	CCA	GAT	AAA	CTT	GAC	ATG	GAA	1006	
Asn	Trp	Asp	Leu	Leu	Asn	Gly	Met	Ser	Pro	Asp	Lys	Leu	Asp	Met	Glu		

- 187 -

320	325	330	335	
GAA GCT TAT GGG CTT ACA TTG CAA CGG GCT TCA CCT TTG ATT GTC CAC				1054
Glu Ala Tyr Gly Leu Thr Leu Gln Arg Ala Ser Pro Leu Ile Val His				
	340	345	350	
CCA AAG CCC AGG CTT GCT AGC TCT ATG TAT GTT AAA T GAAATTATGC				1101
Pro Lys Pro Arg Leu Ala Ser Ser Met Tyr Val Lys				
	355	360		
TGTGCGAATA ATTCCTTATT TATAGCAGGA AATGTCATCT TGAATTATGT GTAATGTTCT				1161
TCTAACTTTC GATGGAAGTG CAAAACAAGT TTTATTAAAA AAAAAAAAAA AAA				1214

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 363 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Arg Ile Leu Thr Arg Ser Ile Ala Ser Ala Gly Glu Asn Pro Ile Asn			
1	5	10	15
Leu Gly Gln Leu Leu Gly Val Cys Thr Thr Asn Ala Leu Ala Arg Val			
20	25	30	
Met Leu Gly Arg Arg Val Phe Gly Asp Gly Ser Gly Gly Val Asp Pro			
35	40	45	
Gln Ala Asp Glu Phe Lys Ser Met Val Val Glu Ile Met Val Leu Ala			
50	55	60	

- 188 -

Gly Ala Phe Asn Leu Gly Asp Phe Ile Pro Ala Leu Asp Trp Phe Asp

65

70

75

80

Leu Gln Gly Ile Thr Ala Lys Met Lys Lys Val His Ala Arg Phe Asp

85

90

95

Ala Phe Leu Asp Ala Ile Leu Glu Glu His Lys Ser Asn Gly Ser Arg

100

105

110

Gly Ala Lys Gln His Val Asp Leu Leu Ser Met Leu Ile Ser Leu Gln

115

120

125

Asp Asn Asn Ile Asp Gly Glu Ser Gly Ala Lys Leu Thr Asp Thr Glu

130

135

140

Ile Lys Ala Leu Leu Leu Asn Leu Phe Thr Ala Gly Thr Asp Thr Ser

145

150

155

160

Ser Ser Thr Val Glu Trp Ala Ile Ala Glu Leu Ile Arg Asn Pro Glu

165

170

175

Val Leu Val Gln Ala Gln Gln Glu Leu Asp Arg Val Val Gly Pro Ser

180

185

190

Arg Leu Val Thr Glu Ser Asp Leu Pro Gln Leu Ala Phe Leu Gln Ala

195

200

205

Val Ile Lys Glu Thr Phe Arg Leu His Pro Ser Thr Pro Leu Ser Leu

210

215

220

Pro Arg Met Ala Ser Glu Gly Cys Glu Ile Asn Gly Tyr Ser Ile Pro

225

230

235

240

Lys Gly Ser Thr Leu Leu Val Asn Val Trp Ser Ile Ala Arg Asp Pro

245

250

255

Ser Ile Trp Ala Asp Pro Leu Glu Phe Arg Pro Ala Arg Phe Leu Pro

- 189 -

260	265	270
Gly Gly Glu Lys Pro Asn Val Asp Val Arg Gly Asn Asp Phe Glu Val		
275	280	285
Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala Gly Met Ser Leu Gly		
290	295	300
Leu Arg Met Val Gln Leu Ser Thr Ala Thr Leu Val His Ser Phe Asn		
305	310	315 320
Trp Asp Leu Leu Asn Gly Met Ser Pro Asp Lys Leu Asp Met Glu Glu		
325	330	335
Ala Tyr Gly Leu Thr Leu Gln Arg Ala Ser Pro Leu Ile Val His Pro		
340	345	350
Lys Pro Arg Leu Ala Ser Ser Met Tyr Val Lys		
355	360	

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1757 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 35..1522

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

- 190 -

CCGTTGCTGT CGAGAAAACA GAAAGAAGAG AAAA ATG GAC TAC GTG AAT ATT	52
Met Asp Tyr Val Asn Ile	
1 5	
TTG CTG GGA CTG TTT TTC ACT TGG TTC TTG GTG AAT GGA CTC ATG TCA	100
Leu Leu Gly Leu Phe Phe Thr Trp Phe Leu Val Asn Gly Leu Met Ser	
10 15 20	
CTT CGA AGA AGA AAA ATC TCT AAG AAA CTT CCA CCA GGT CCA TTT CCT	148
Leu Arg Arg Arg Lys Ile Ser Lys Lys Leu Pro Pro Gly Pro Phe Pro	
25 30 35	
TTG CCT ATC ATC GGA AAT CTT CAC TTA CTT GGT AAT CAT CCT CAC AAA	196
Leu Pro Ile Ile Gly Asn Leu His Leu Leu Gly Asn His Pro His Lys	
40 45 50	
TCA CTT GCT CAA CTT GCA AAA ATT CAT GGT CCT ATT ATG AAT CTC AAA	244
Ser Leu Ala Gln Leu Ala Lys Ile His Gly Pro Ile Met Asn Leu Lys	
55 60 65 70	
TTA GGC CAA CTA AAC ACA GTG GTC ATT TCA TCA TCA GTC GTG GCA AGA	292
Leu Gly Gln Leu Asn Thr Val Val Ile Ser Ser Ser Val Val Ala Arg	
75 80 85	
GAA GTC TTG CAA AAA CAA GAC TTA ACA TTT TCC AAT AGG TTT GTC CCG	340
Glu Val Leu Gln Lys Gln Asp Leu Thr Phe Ser Asn Arg Phe Val Pro	
90 95 100	
GAC GTA GTC CAT GTC CGA AAT CAC TCC GAT TTT TCT GTT GTT TGG TTA	388
Asp Val Val His Val Arg Asn His Ser Asp Phe Ser Val Val Trp Leu	
105 110 115	
CCA GTC AAT TCT CGA TGG AAA ACG CTT CGC AAA ATC ATG AAC TCT AGC	436
Pro Val Asn Ser Arg Trp Lys Thr Leu Arg Lys Ile Met Asn Ser Ser	
120 125 130	
ATC TTT TCT GGT AAC AAG CTT GAT GGT AAT CAA CAT CTG AGG TCT AAA	484

- 191 -

Ile	Phe	Ser	Gly	Asn	Lys	Leu	Asp	Gly	Asn	Gln	His	Leu	Arg	Ser	Lys	
135					140					145					150	
AAG	GTC	CAA	GAG	TTA	ATT	GAT	TAT	TGT	CAA	AAG	TGT	GCC	AAG	AAT	GGC	532
Lys	Val	Gln	Glu	Leu	Ile	Asp	Tyr	Cys	Gln	Lys	Cys	Ala	Lys	Asn	Gly	
				155					160					165		
GAA	GCA	GTG	GAT	ATA	GGA	AGA	GCA	ACT	TTT	GGA	ACT	ACT	TTG	AAT	TTG	580
Glu	Ala	Val	Asp	Ile	Gly	Arg	Ala	Thr	Phe	Gly	Thr	Thr	Leu	Asn	Leu	
				170				175					180			
CTA	TCC	AAC	ACC	ATT	TTC	TCT	AAA	GAT	TTG	ACT	AAT	CCG	TTT	TCT	GAT	628
Leu	Ser	Asn	Thr	Ile	Phe	Ser	Lys	Asp	Leu	Thr	Asn	Pro	Phe	Ser	Asp	
				185				190					195			
TCT	GCT	AAA	GAG	TTT	AAG	GAA	TTG	GTT	TGG	AAC	ATT	ATG	GTT	GAG	GCT	676
Ser	Ala	Lys	Glu	Phe	Lys	Glu	Leu	Val	Trp	Asn	Ile	Met	Val	Glu	Ala	
				200				205				210				
GGA	AAA	CCC	AAT	TTG	GTG	GAC	TAC	TTT	CCT	TTC	CTT	GAG	AAA	ATT	GAT	724
Gly	Lys	Pro	Asn	Leu	Val	Asp	Tyr	Phe	Pro	Phe	Leu	Glu	Lys	Ile	Asp	
215					220					225				230		
CCG	CAA	GGT	ATA	AAG	CGA	CGC	ATG	ACT	AAT	AAT	TTT	ACT	AAG	TTT	CTT	772
Pro	Gln	Gly	Ile	Lys	Arg	Arg	Met	Thr	Asn	Asn	Phe	Thr	Lys	Phe	Leu	
				235					240					245		
GGC	CTT	ATC	AGC	GGT	TTG	ATT	GAT	GAC	CGG	TTA	AAG	GAA	AGG	AAT	CTA	820
Gly	Leu	Ile	Ser	Gly	Leu	Ile	Asp	Asp	Arg	Leu	Lys	Glu	Arg	Asn	Leu	
				250					255					260		
AGG	GAC	AAT	GCA	AAT	ATT	GAT	GTT	TTA	GAC	GCC	CTT	CTC	AAC	ATT	AGC	868
Arg	Asp	Asn	Ala	Asn	Ile	Asp	Val	Leu	Asp	Ala	Leu	Leu	Asn	Ile	Ser	
				265				270						275		
CAA	GAG	AAC	CCA	GAA	GAG	ATT	GAC	AGG	AAT	CAA	ATC	GAG	CAG	TTG	TGT	916
Gln	Glu	Asn	Pro	Glu	Glu	Ile	Asp	Arg	Asn	Gln	Ile	Glu	Gln	Leu	Cys	

- 192 -

280	285	290	
CTG GAC TTG TTT GCA GCA GGG ACT GAT ACT ACA TCG AAT ACC TTG GAG			964
Leu Asp Leu Phe Ala Ala Gly Thr Asp Thr Thr Ser Asn Thr Leu Glu			
295	300	305	310
TGG GCA ATG GCA GAA CTA CTT CAG AAT CCA CAC ACA TTG CAG AAA GCA			1012
Trp Ala Met Ala Glu Leu Leu Gln Asn Pro His Thr Leu Gln Lys Ala			
315	320	325	
CAA GAA GAA CTT GCA CAA GTC ATT GGT AAA GGC AAA CAA GTA GAA GAA			1060
Gln Glu Glu Leu Ala Gln Val Ile Gly Lys Gly Lys Gln Val Glu Glu			
330	335	340	
GCA GAT GTT GGA CGA CTA CCT TAC TTG CGA TGC ATA GTG AAA GAA ACC			1108
Ala Asp Val Gly Arg Leu Pro Tyr Leu Arg Cys Ile Val Lys Glu Thr			
345	350	355	
TTA CGA ATA CAC CCA GCG GCT CCT CTC TTA ATT CCA CGT AAA GTG GAG			1156
Leu Arg Ile His Pro Ala Ala Pro Leu Leu Ile Pro Arg Lys Val Glu			
360	365	370	
GAA GAC GTT GAG TTG TCT ACC TAT ATT ATT CCA AAG GAT TCA CAA GTT			1204
Glu Asp Val Glu Leu Ser Thr Tyr Ile Ile Pro Lys Asp Ser Gln Val			
375	380	385	390
CTA GTG AAC GTA TGG GCA ATT GGA CGC AAC TCT GAT CTA TGG GAA AAT			1252
Leu Val Asn Val Trp Ala Ile Gly Arg Asn Ser Asp Leu Trp Glu Asn			
395	400	405	
CCT TTG GTC TTT AAG CCA GAA AGG TTT TGG GAG TCA GAA ATA GAT ATC			1300
Pro Leu Val Phe Lys Pro Glu Arg Phe Trp Glu Ser Glu Ile Asp Ile			
410	415	420	
CGA GGT CGA GAT TTT GAA CTC ATT CCA TTT GGT GCT GGT CGA AGA ATT			1348
Arg Gly Arg Asp Phe Glu Leu Ile Pro Phe Gly Ala Gly Arg Arg Ile			
425	430	435	

- 193 -

TGC CCT GGA TTG CCT TTG GCT ATG AGG ATG ATT CCA GTA GCA CTA GGT	1396
Cys Pro Gly Leu Pro Leu Ala Met Arg Met Ile Pro Val Ala Leu Gly	
440 445 450	
 TCA TTG CTA AAC TCA TTT AAT TGG AAA CTA TAT GGT GGA ATT GCA CCT	1444
Ser Leu Leu Asn Ser Phe Asn Trp Lys Leu Tyr Gly Gly Ile Ala Pro	
455 460 465 470	
 AAA GAT TTG GAC ATG CAG GAA AAG TTT GGC ATT ACC TTG GCG AAA GCC	1492
Lys Asp Leu Asp Met Gln Glu Lys Phe Gly Ile Thr Leu Ala Lys Ala	
475 480 485	
 CAA CCT CTG CTA GCT ATC CCA ACT CCC CTG TAGCTATAGG GATAAATTAA	1542
Gln Pro Leu Leu Ala Ile Pro Thr Pro Leu	
490 495	
 GTTGAGGTTT TAAGTTACTA GTAGATTCTA TTGCAGCTAT AGGATTTCTT TCACCATCAC	1602
 GTATGCTTTA CCGTTGGATG ATGGAAAGAA ATATCTATAG CTTTGGGTTT GTTTAGTTTG	1662
 CACATAAAAA TTGAATGAAT GGAATACCAT GGAGTTATAA GAAATAATAA GACTATGATT	1722
 CTTACCCTAC TTGAACAATG ACATGGCTAT TTCAC	1757

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

- 194 -

TTTTTTTTTT TTTTTTA

18

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TTTTTTTTTT TTTTTTC

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TTTTTTTTTT TTTTTTG

18

- 195 -

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Trp Ala Ile Gly Arg Asp Pro

5

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TGGGCIATIG GI (A/C) GIGA(T/C) CC

- 196 -

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Phe Arg Pro Glu Arg Phe

5

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 acids
- (B) TYPE: nucleic acids
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AGGAATT(T/C)(A/C)G ICCIGA(A/G)(A/C)GI TT

22

- 197 -

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 nucleic acids
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CCITT(T/C)GGIG CIGGI(A/C)GI(A/C)G IATITG(T/G)(C/G)CI GG

32

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Glu Phe Xaa Pro Glu Arg Phe

- 198 -

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 nucleic acids
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GAITT(T/C)IIIC CIGAI(A/C)GITT

20

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 nucleic acids
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CCACACGAGT AGTTTTGGCA TTTGACCC

28

- 199 -

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 nucleic acids
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GTCTTGGACA TCACACTTCA ATCTG

25

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 nucleic acids
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

CCGAATTCCC CCCCCC

17

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 nucleic acids

- 200 -

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

CCIGG(A/G)CAIA TIC(G/T) (C/T) (C/T)TICC IGCICC(A/G)AAI GG

32

- 201 -

CLAIMS:

1. An isolated nucleic acid molecule comprising a sequence of nucleotides encoding a flavonoid 3'-hydroxylase or a derivative thereof wherein said flavonoid 3'-hydroxylase or its derivative is capable of more efficient modulation of hydroxylation of flavonoid compounds in plants than is a flavonoid 3'-hydroxylase encoded by the nucleotide sequence set forth in SEQ ID NO:26.
2. An isolated nucleic acid molecule according to claim 1 comprising a nucleotide sequence which corresponds to the genetic locus designated *Ht1* or *Ht2* in petunia or to loci in other flowering plant species which contain sequences which control production of 3'-hydroxylated flavonoids.
3. An isolated nucleic acid molecule according to claim 2 comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:1 or having at least about 60% similarity thereto or capable of hybridizing to the sequence set forth in SEQ ID NO:1 under low stringency conditions.
4. An isolated nucleic acid molecule according to claim 2 comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:3 or having at least about 60% similarity thereto or capable of hybridizing to the sequence set forth in SEQ ID NO:3 under low stringency conditions.
5. An isolated nucleic acid molecule according to claim 2 comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:5 or having at least about 60% similarity thereto or capable of hybridizing to the sequence set forth in SEQ ID NO:5 under low stringency conditions.
6. An isolated nucleic acid molecule according to claim 2 comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:7

- 202 -

or having at least about 60% similarity thereto or capable of hybridizing to the sequence set forth in SEQ ID NO:7 under low stringency conditions.

7. An isolated nucleic acid molecule according to claim 2 comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:9 or having at least about 60% similarity to the coding region thereof or capable of hybridizing to the sequence set forth in SEQ ID NO:9 under low stringency conditions.

8. An isolated nucleic acid molecule according to claim 2 comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:14 or having at least about 60% similarity thereto or capable of hybridizing to the sequence set forth in SEQ ID NO:14 under low stringency conditions.

9. An isolated nucleic acid molecule according to claim 2 comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:16 or having at least about 60% similarity thereto or capable of hybridizing to the sequence set forth in SEQ ID NO:16 under low stringency conditions.

10. An isolated nucleic acid molecule according to claim 2 comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:18 or having at least about 60% similarity thereto or capable of hybridizing to the sequence set forth in SEQ ID NO:18 under low stringency conditions.

11. An isolated nucleic acid molecule according to claim 2 comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:20 or having at least about 60% similarity thereto or capable of hybridizing to the sequence set forth in SEQ ID NO:20 under low stringency conditions.

12. An isolated nucleic acid molecule according to claim 2 comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:22

- 203 -

or having at least about 60% similarity thereto or capable of hybridizing to the sequence set forth in SEQ ID NO:22 under low stringency conditions.

13. An isolated nucleic acid molecule according to claim 2 comprising a nucleotide sequence or complementary nucleotide sequence substantially as set forth in SEQ ID NO:24 or having at least about 60% similarity thereto or capable of hybridizing to the sequence set forth in SEQ ID NO:24 under low stringency conditions.

14. An isolated nucleic acid molecule according to claim 2 comprising a sequence of nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:2 or an amino acid sequence having at least about 50% similarity thereto.

15. An isolated nucleic acid molecule according to claim 2 comprising a sequence of nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:4 or an amino acid sequence having at least about 50% similarity thereto.

16. An isolated nucleic acid molecule according to claim 2 comprising a sequence of nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:6 or an amino acid sequence having at least about 50% similarity thereto.

17. An isolated nucleic acid molecule according to claim 2 comprising a sequence of nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:8 or an amino acid sequence having at least about 50% similarity thereto.

18. An isolated nucleic acid molecule according to claim 2 comprising a sequence of nucleotides encoding or complementary to a sequence encoding an amino acid sequence

- 204 -

substantially as set forth in SEQ ID NO:10 or SEQ ID NO:11 or SEQ ID NO:12 or SEQ ID NO:13 or an amino acid sequence having at least about 50% similarity thereto.

19. An isolated nucleic acid molecule according to claim 2 comprising a sequence of nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:15 or an amino acid sequence having at least about 50% similarity thereto.

20. An isolated nucleic acid molecule according to claim 2 comprising a sequence of nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:17 or an amino acid sequence having at least about 50% similarity thereto.

21. An isolated nucleic acid molecule according to claim 2 comprising a sequence of nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:19 or an amino acid sequence having at least about 50% similarity thereto.

22. An isolated nucleic acid molecule according to claim 2 comprising a sequence of nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:21 or an amino acid sequence having at least about 50% similarity thereto.

23. An isolated nucleic acid molecule according to claim 2 comprising a sequence of nucleotides encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:23 or an amino acid sequence having at least about 50% similarity thereto.

24. An isolated nucleic acid molecule according to claim 2 comprising a sequence of nucleotides encoding or complementary to a sequence encoding an amino acid sequence

- 205 -

substantially as set forth in SEQ ID NO:25 or an amino acid sequence having at least about 50% similarity thereto.

25. An oligonucleotide capable of hybridizing under low stringency conditions to a nucleotide sequence selected from SEQ ID NO:1, 3, 5, 7, 9, 14, 16, 18, 20, 22 and 24.

26. A genetic construct capable of reducing expression of an endogenous gene encoding a flavonoid 3'-hydroxylase in a plant, said genetic construct comprising a nucleotide sequence selected from:

- (i) a nucleotide sequence encoding an amino acid sequence set forth in one of SEQ ID NO:2, 4, 6, 8, 10, 11, 12, 13, 15, 17, 19, 21, 23 or 25 or a complementary form thereof;
- (ii) a nucleotide sequence set forth in one of SEQ ID NO:1, 3, 5, 7, 14, 16, 18, 20, 22 or 24 or the coding region in SEQ ID NO:9 or a complementary form thereof;
- (iii) a nucleotide sequence having at least about 60% similarity to (i) or (ii); and
- (iv) a nucleotide sequence capable of hybridizing under low stringency conditions to (i), (ii) and/or (iii).

27. A method for producing a transgenic plant capable of synthesizing a flavonoid 3'-hydroxylase or a functional derivative thereof, said method comprising stably transforming a cell of a suitable plant with nucleic acid molecule which comprises a sequence of nucleotides encoding said flavonoid 3'-hydroxylase or a derivative thereof under conditions permitting the eventual expression of said nucleic acid molecule, regenerating a transgenic plant from the cell and growing said transgenic plant for a time and under conditions sufficient to permit the expression of the nucleic acid molecule.

28. A method for producing a transgenic plant with reduced endogenous or existing flavonoid 3'-hydroxylase activity, said method comprising stably transforming a cell of a suitable plant with a nucleic acid molecule which comprises a sequence of nucleotides encoding or complementary to a sequence encoding flavonoid 3'-hydroxylase or a derivative thereof, regenerating a transgenic plant from the cell and where necessary growing said

- 206 -

transgenic plant under conditions sufficient to permit the expression of the nucleic acid molecule.

29. A method according to claim 27 or 28 wherein the introduced nucleic acid molecule comprises a nucleotide sequence or complementary nucleotide sequence selected from:

- (i) a nucleotide sequence encoding an amino acid sequence set forth in one of SEQ ID NO:2, 4, 6, 8, 10, 11, 12, 13, 15, 17, 19, 21, 23 or 25 or a complementary form thereof;
- (ii) a nucleotide sequence set forth in one of SEQ ID NO:1, 3, 5, 7, 14, 16, 18, 20, 22 or 24 or the coding region in SEQ ID NO:9 or a complementary form thereof;
- (iii) a nucleotide sequence having at least about 60% similarity to (i) or (ii); and
- (iv) a nucleotide sequence capable of hybridizing under low stringency conditions to (i), (ii) and/or (iii).

30. A method according to claim 27 or 28 wherein the recipient plant is selected from petunia, carnation, chrysanthemum, rose, snapdragon, tobacco, cornflower, pelargonium, lisianthus, gerbera, apple, iris, lily, African violet and morning glory.

31. A method for producing a transgenic plant capable of modulating hydroxylation of flavonoid compounds, said method comprising stably transforming a cell or group of cells of a suitable plant with a nucleic acid molecule comprising a sequence of nucleotides encoding, or complementary to a sequence encoding, flavonoid 3'-hydroxylase or a derivative thereof, and regenerating a transgenic plant from said cell or cells.

32. A method according to claim 31 where the transformed nucleic acid molecule comprises a nucleotide sequence selected from:

- (i) a nucleotide sequence encoding an amino acid sequence set forth in one of SEQ ID NO:2, 4, 6, 8, 10, 11, 12, 13, 15, 17, 19, 21, 23 or 25 or a complementary form thereof;
- (ii) a nucleotide sequence set forth in one of SEQ ID NO:1, 3, 5, 7, 14, 16, 18, 20, 22 or 24 or the coding region in SEQ ID NO:9 or a complementary form thereof;
- (iii) a nucleotide sequence having at least about 60% similarity to (i) or (ii); and

- 207 -

(iv) a nucleotide sequence capable of hybridizing under low stringency conditions to (i), (ii) and/or (iii).

33. A transgenic plant having tissue exhibiting altered colour, said transgenic plant comprising a nucleic acid molecule comprising a sequence of nucleotides selected from:

(i) a nucleotide sequence encoding an amino acid sequence set forth in one of SEQ ID NO:2, 4, 6, 8, 10, 11, 12, 13, 15, 17, 19, 21, 23 or 25 or a complementary form thereof;

(ii) a nucleotide sequence set forth in one of SEQ ID NO:1, 3, 5, 7, 14, 16, 18, 20, 22 or 24 or the coding region in SEQ ID NO:9 or a complementary form thereof;

(iii) a nucleotide sequence having at least about 60% similarity to (i) or (ii); and

(iv) a nucleotide sequence capable of hybridizing under low stringency conditions to (i), (ii) and/or (iii).

34. A cut flower from a transgenic plant according to claim 33.

35. A seed from a transgenic plant according to claim 33.

36. A fruit from a transgenic plant according to claim 33.

37. A leaf from a transgenic plant according to claim 33.

38. Use of a nucleic acid molecule comprising a sequence of nucleotides encoding a flavonoid 3'-hydroxylase in the manufacture of a genetic construct capable of modulating hydroxylation of flavonoid compounds in a plant or cells of a plant.

39. Use according to claim 38 wherein the nucleotide sequence is selected from:

(i) a nucleotide sequence encoding an amino acid sequence set forth in one of SEQ ID NO:2, 4, 6, 8, 10, 11, 12, 13, 15, 17, 19, 21, 23 or 25 or a complementary form thereof;

(ii) a nucleotide sequence set forth in one of SEQ ID NO:1, 3, 5, 7, 14, 16, 18, 20, 22 or 24 or the coding region in SEQ ID NO:9 or a complementary form thereof;

- 208 -

- (iii) a nucleotide sequence having at least about 60% similarity to (i) or (ii); and
- (iv) a nucleotide sequence capable of hybridizing under low stringency conditions to (i), (ii) and/or (iii).

1/21



Figure 1a



Figure 1b

2/21

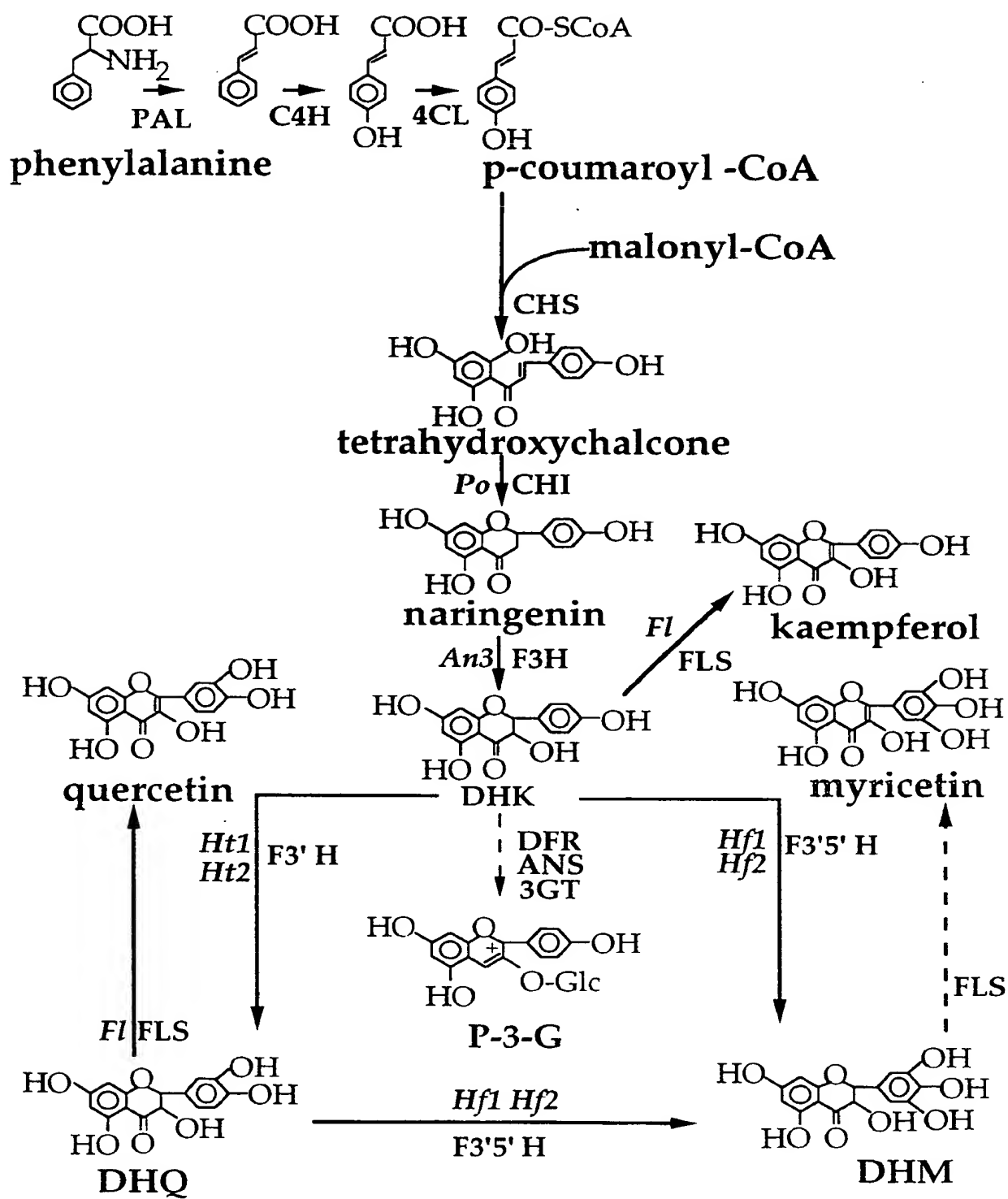


Figure 1a

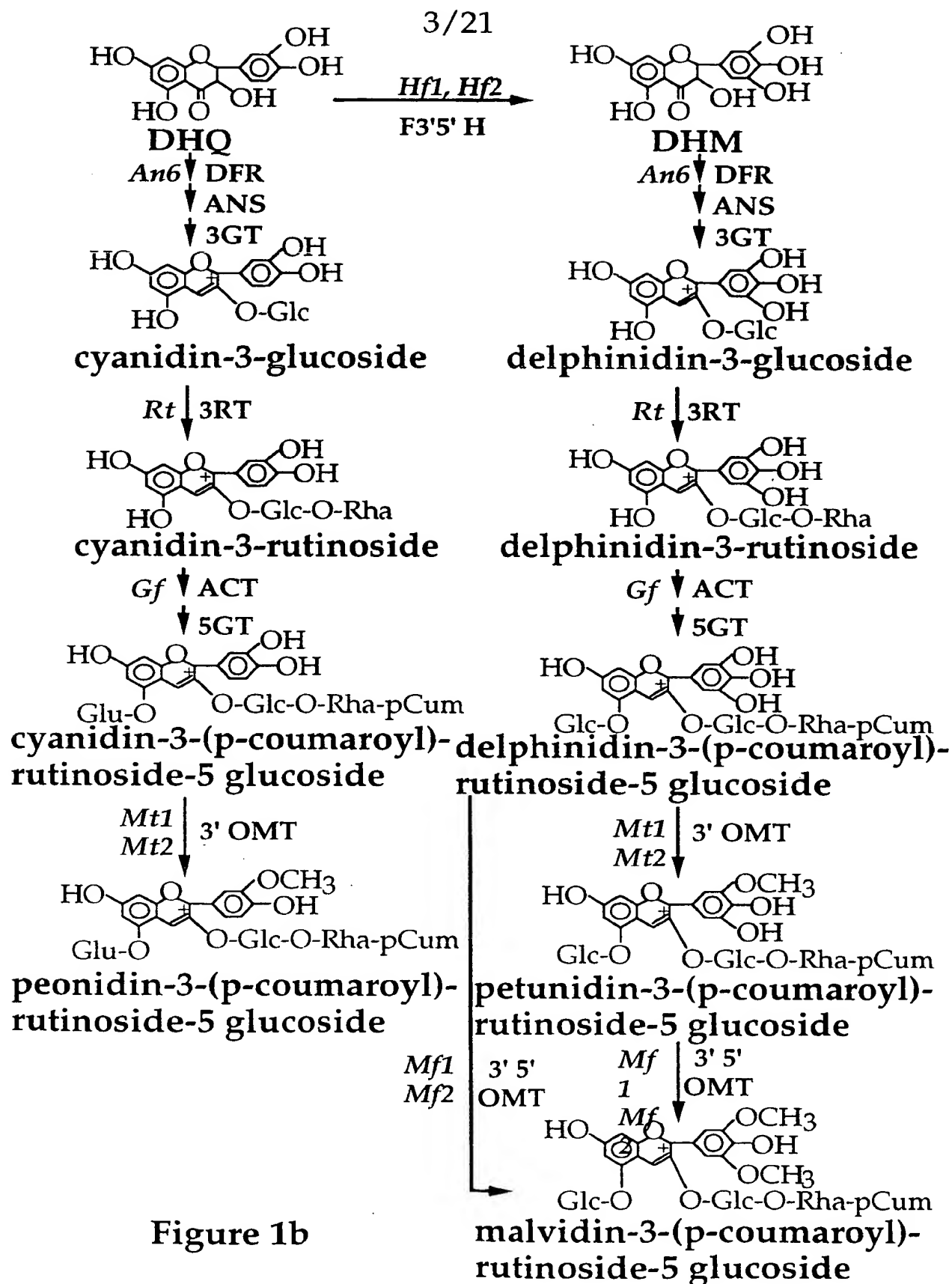


Figure 1b

4/21

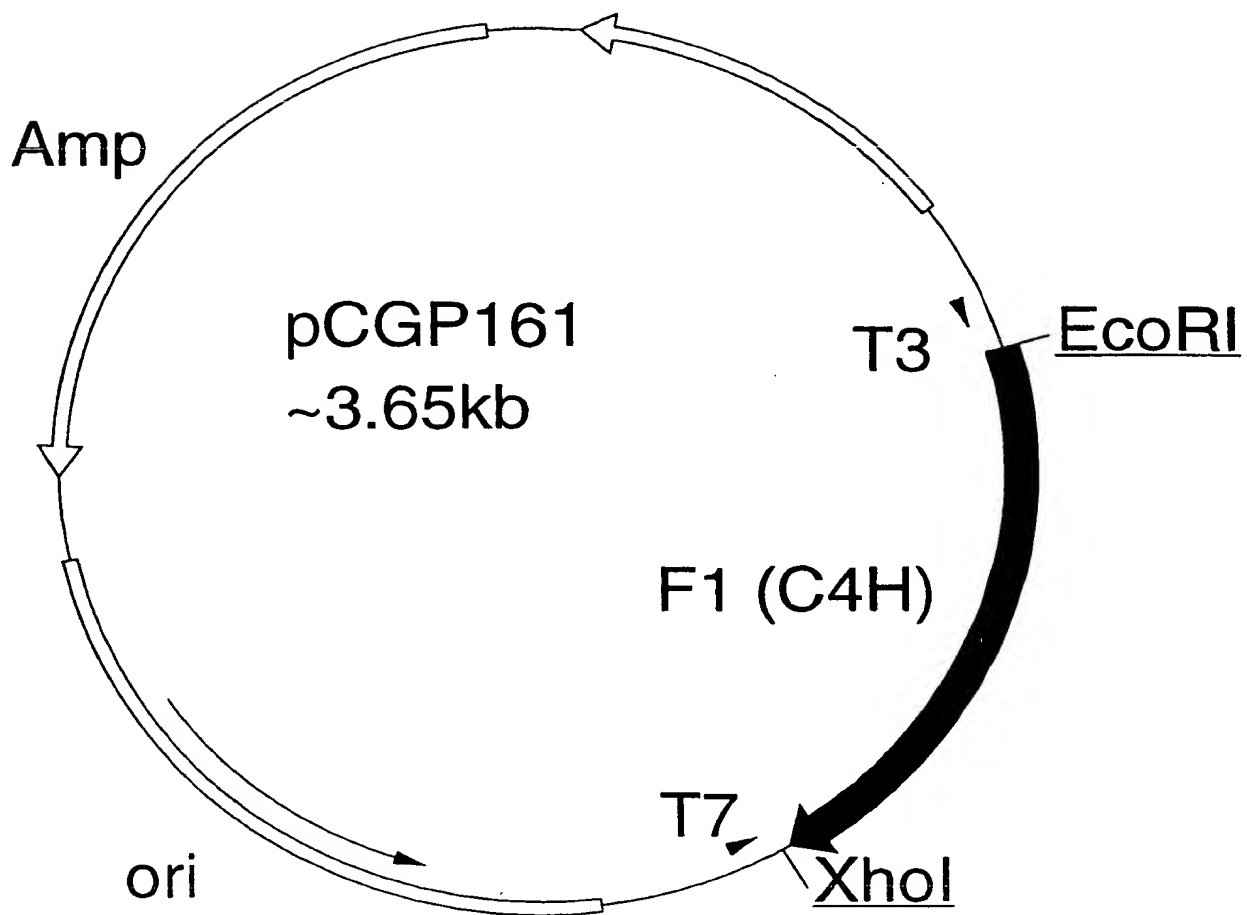


Figure 2

5/21

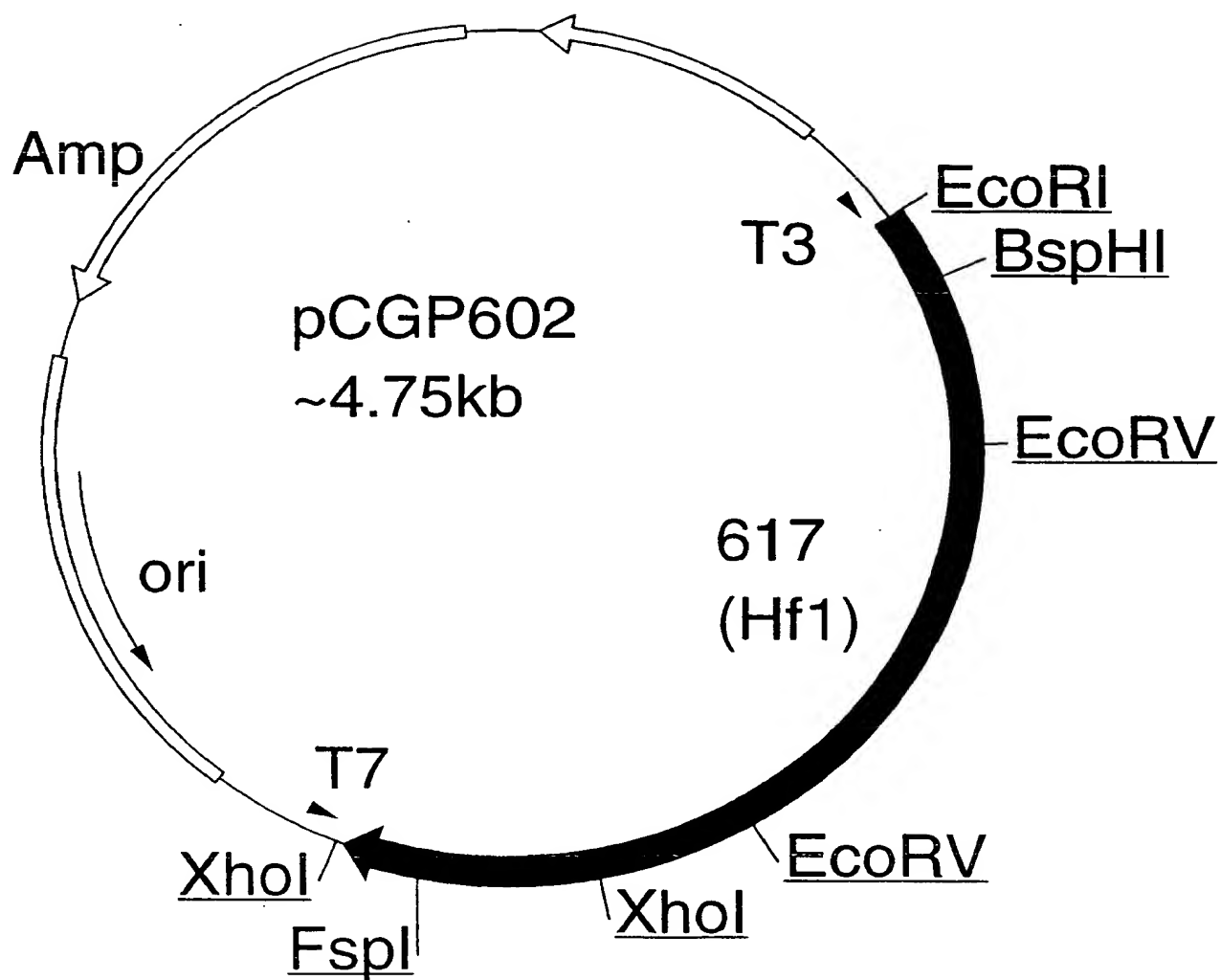


Figure 3

6/21

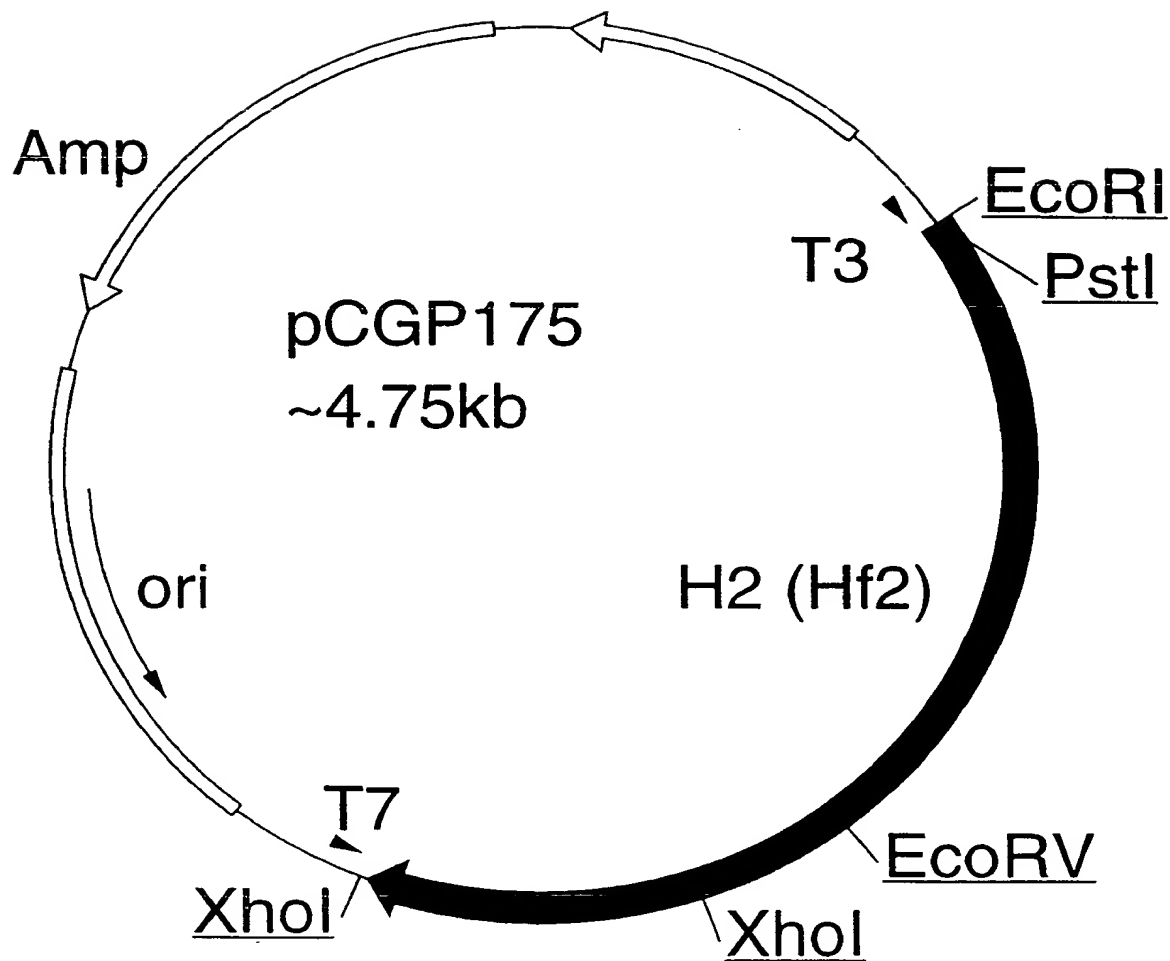


Figure 4

7/21

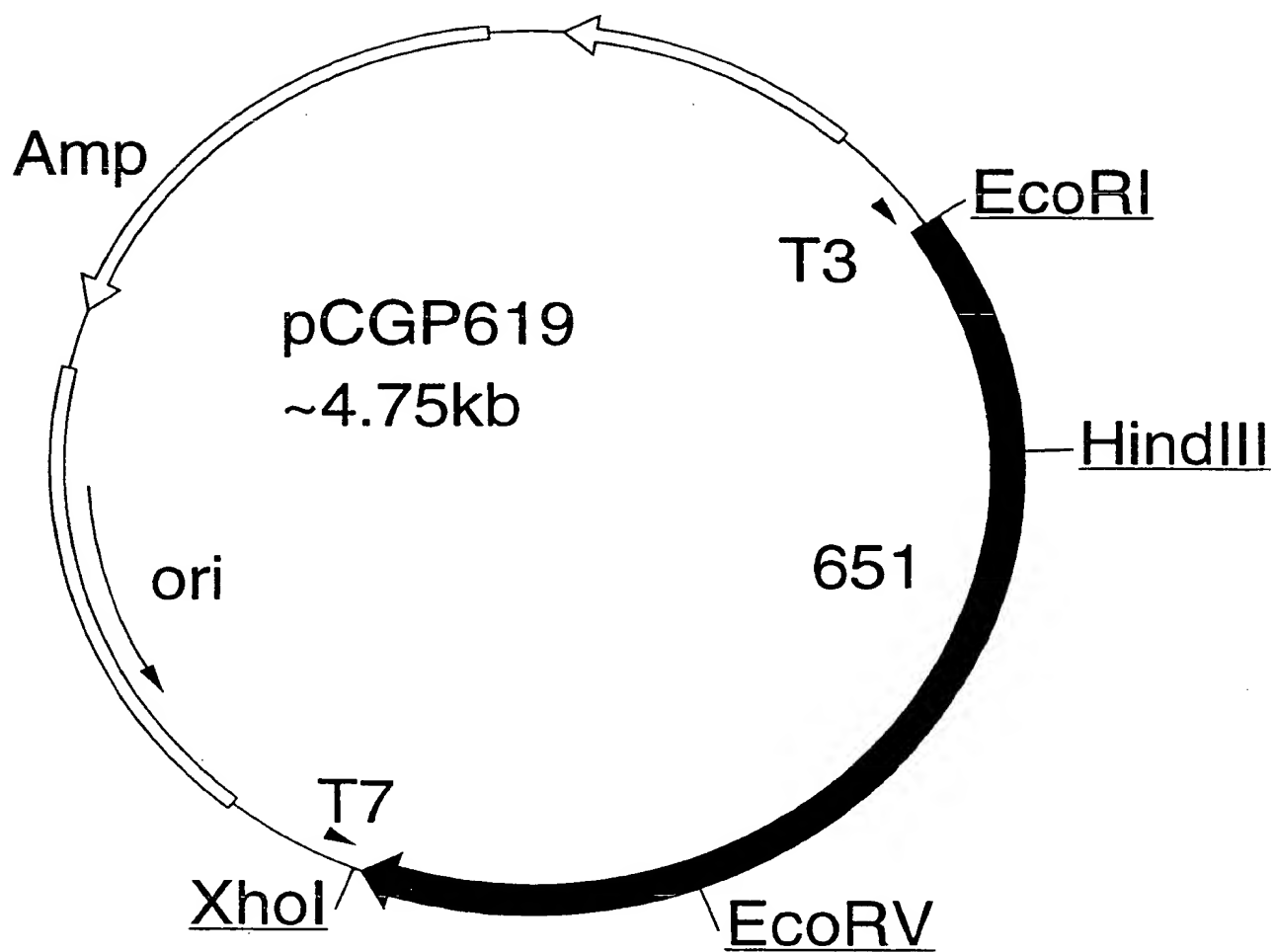
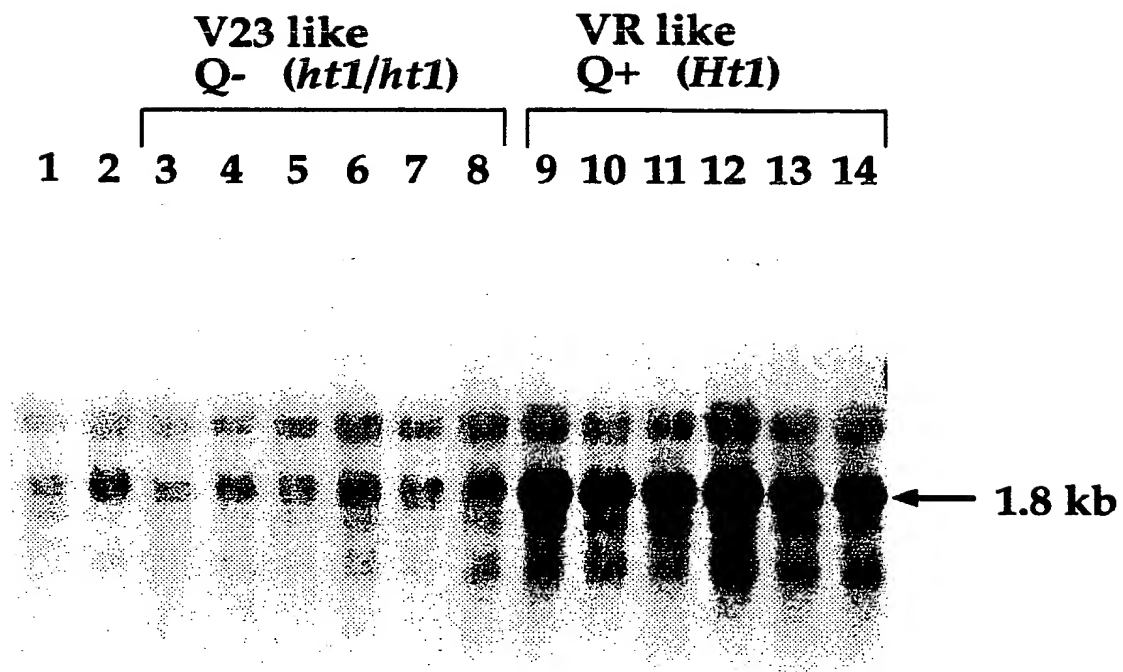


Figure 5

8/21

**Figure 6**

9/21

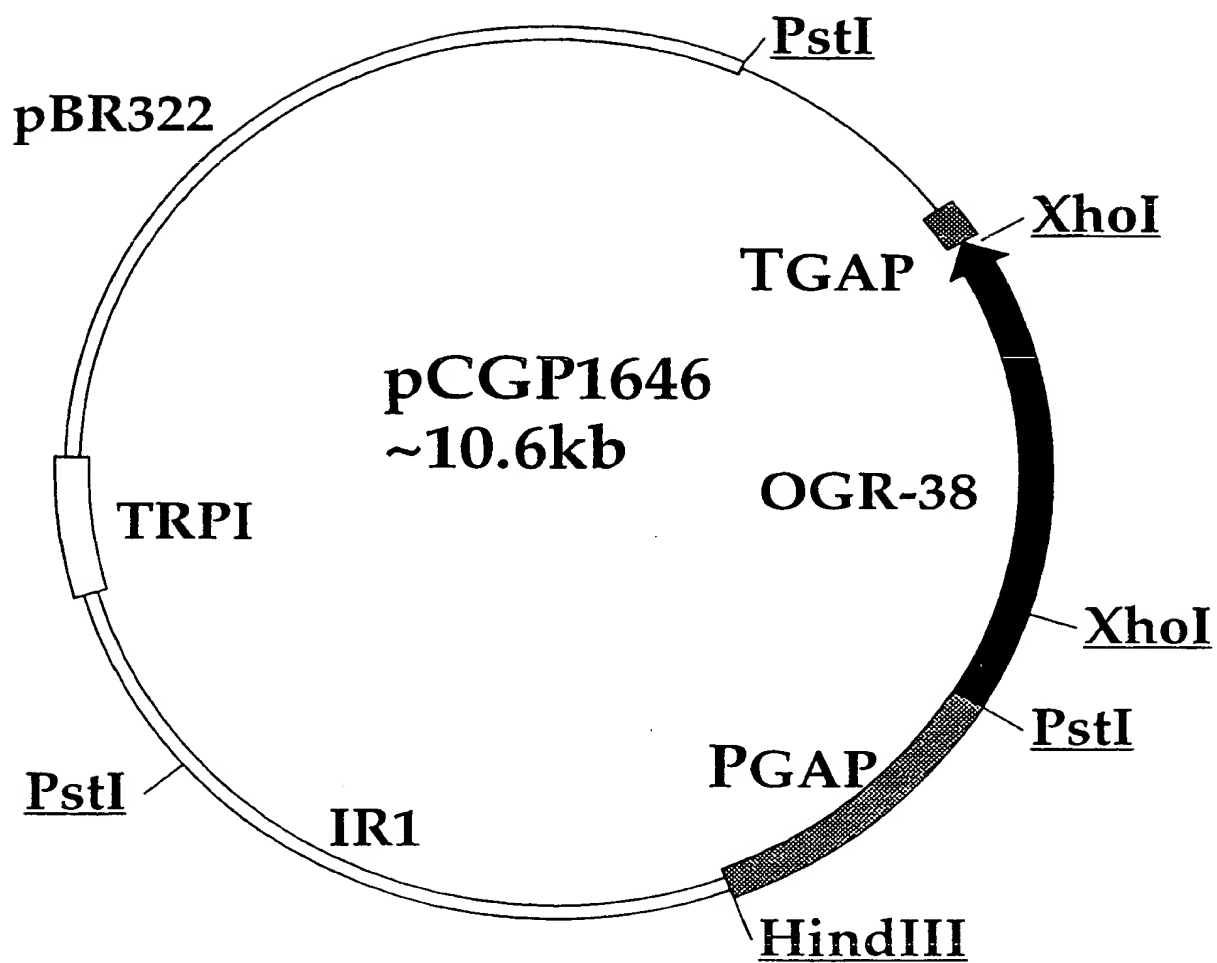


Figure 7

10/21

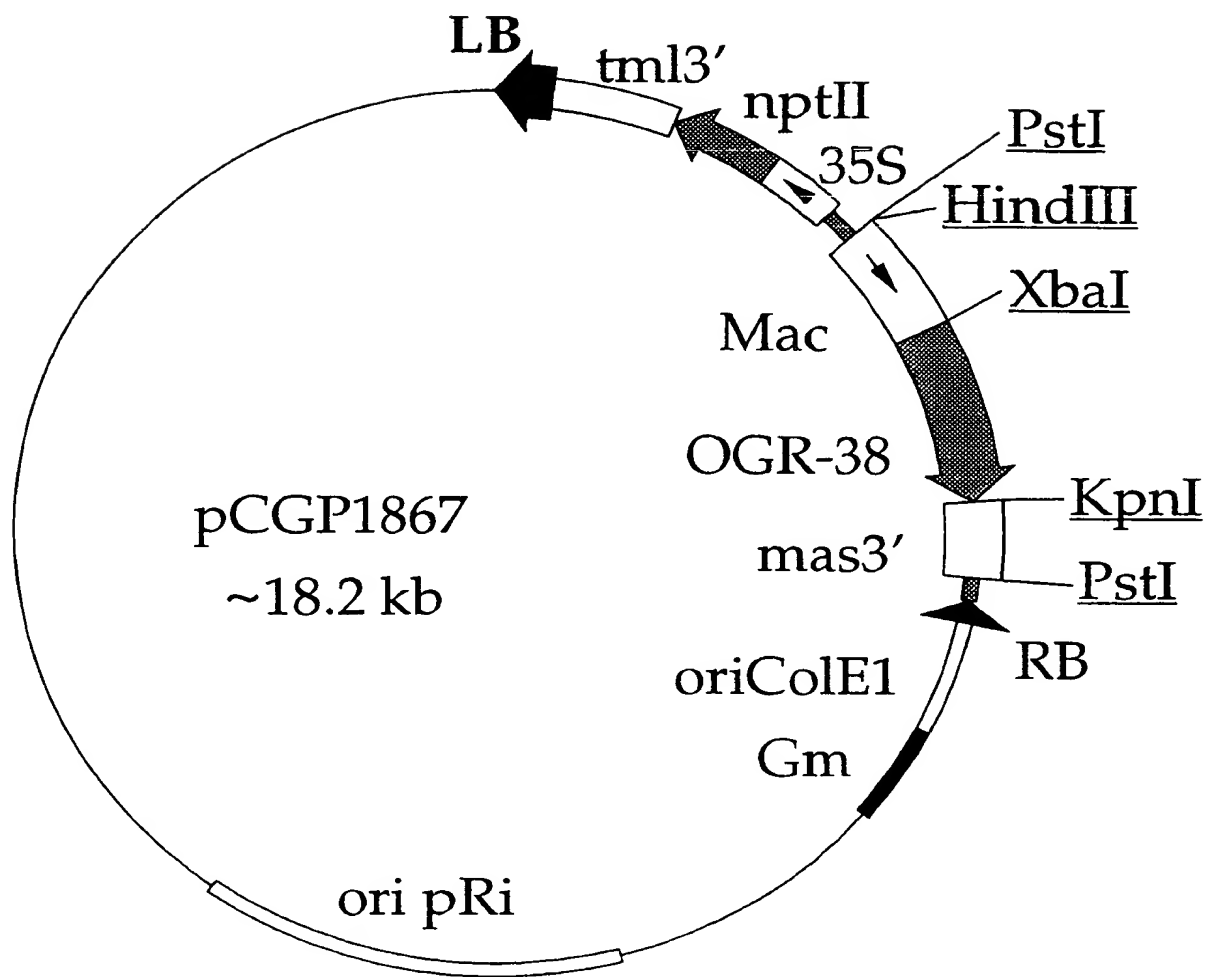


Figure 8

11/21

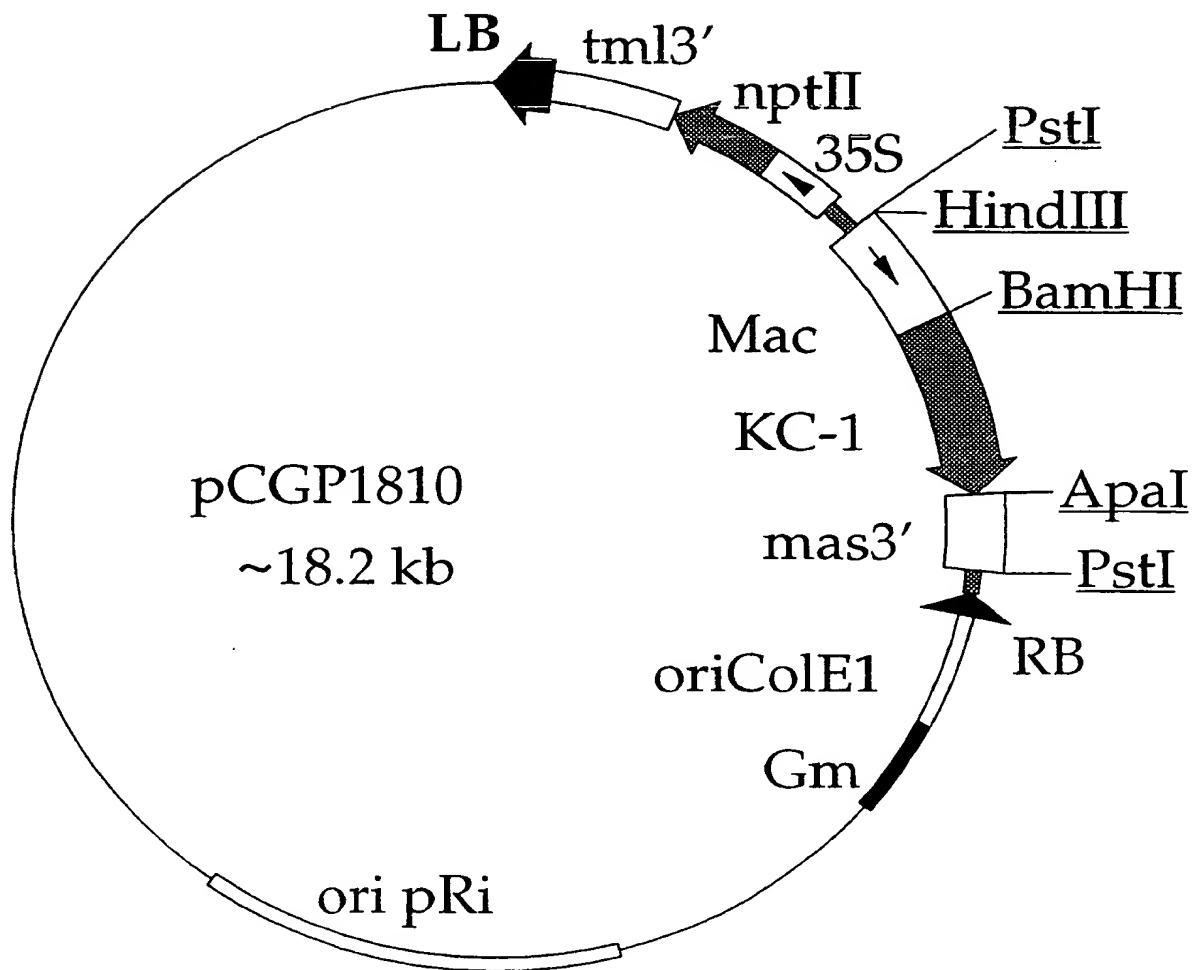


Figure 9

12/21

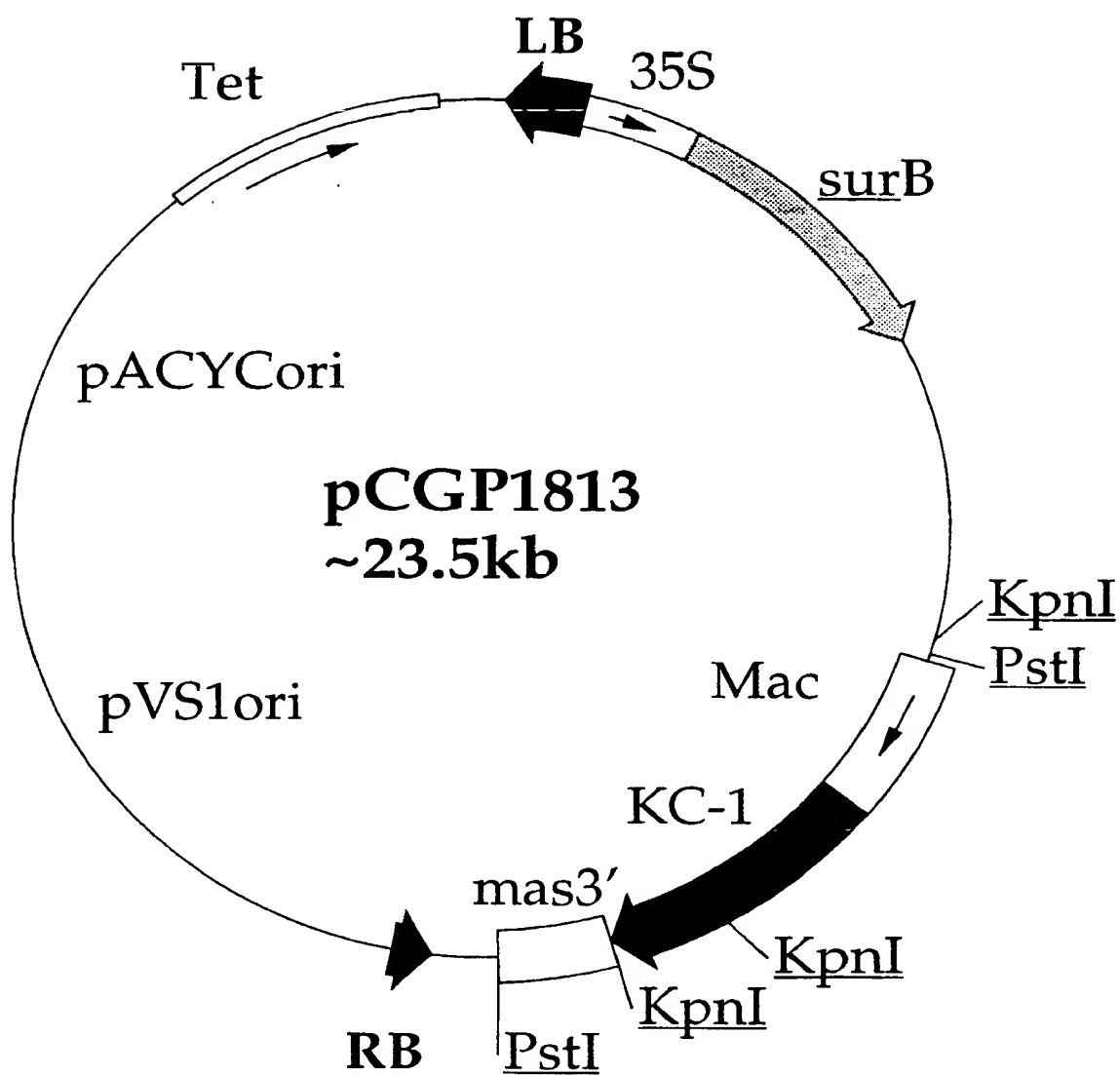


Figure 10

13/21

N8K16		N8 x K16 F2 population												
+	-	+	+	+	+	+	-	+	+	-	+	-	-	+
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

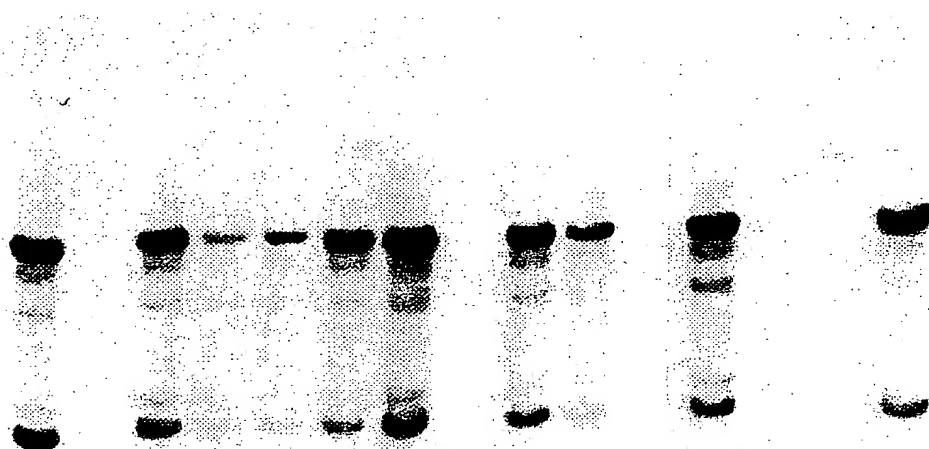
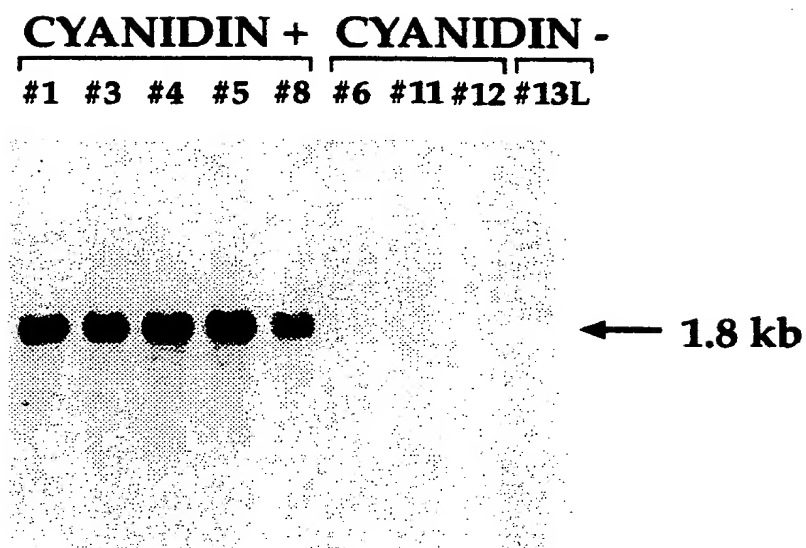


Figure 11

14/21

**Figure 12**

15/21

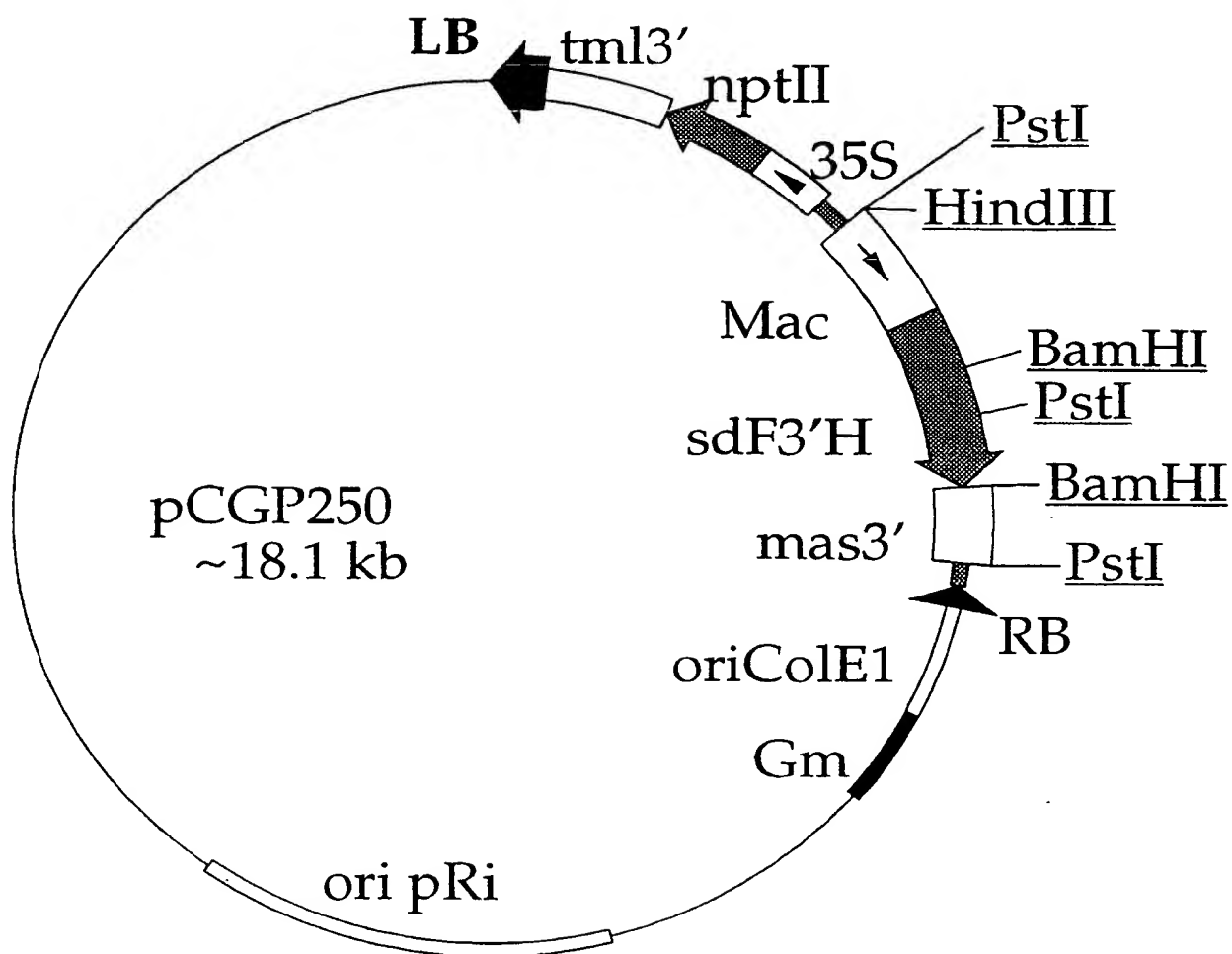


Figure 13

16/21

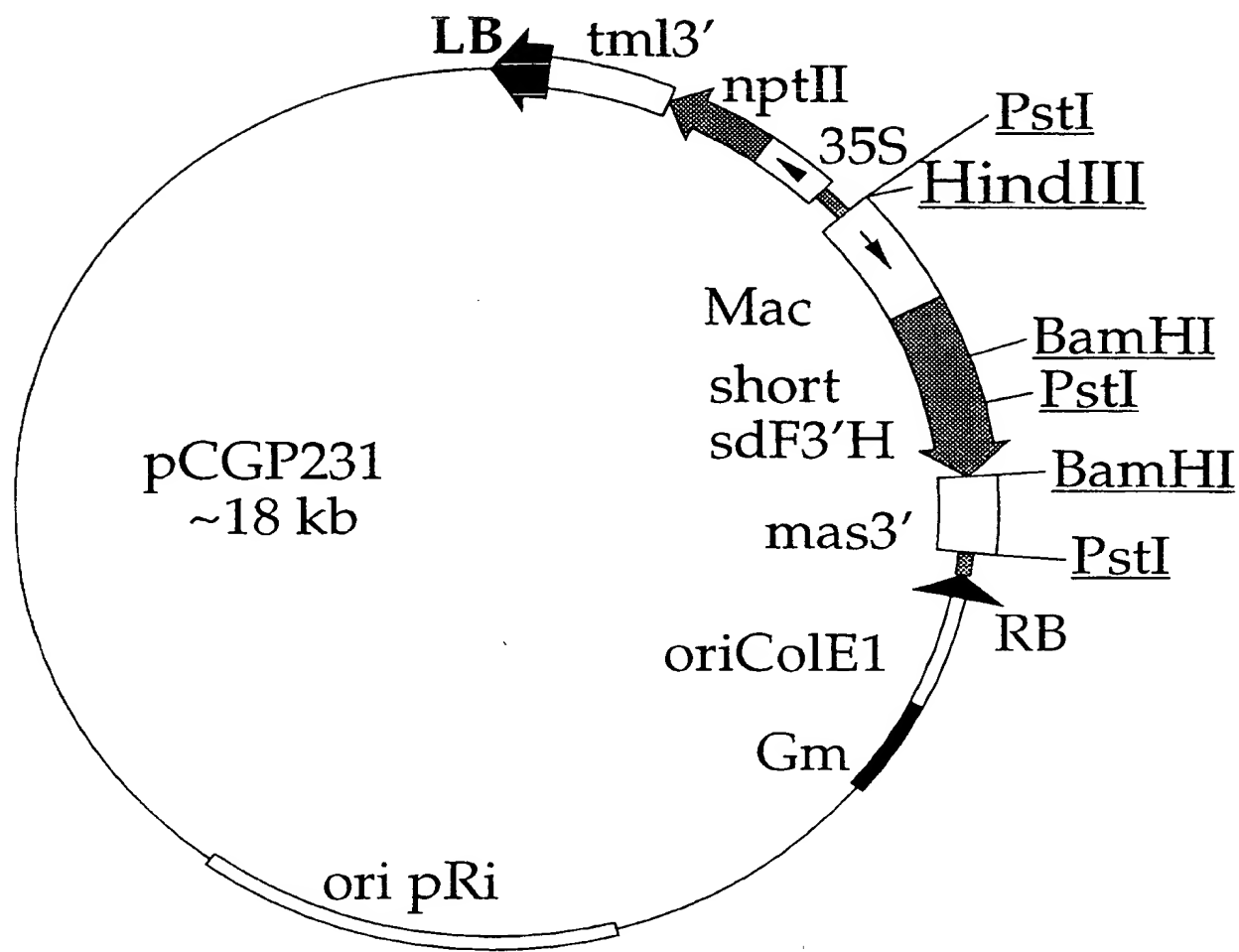


Figure 14

17/21

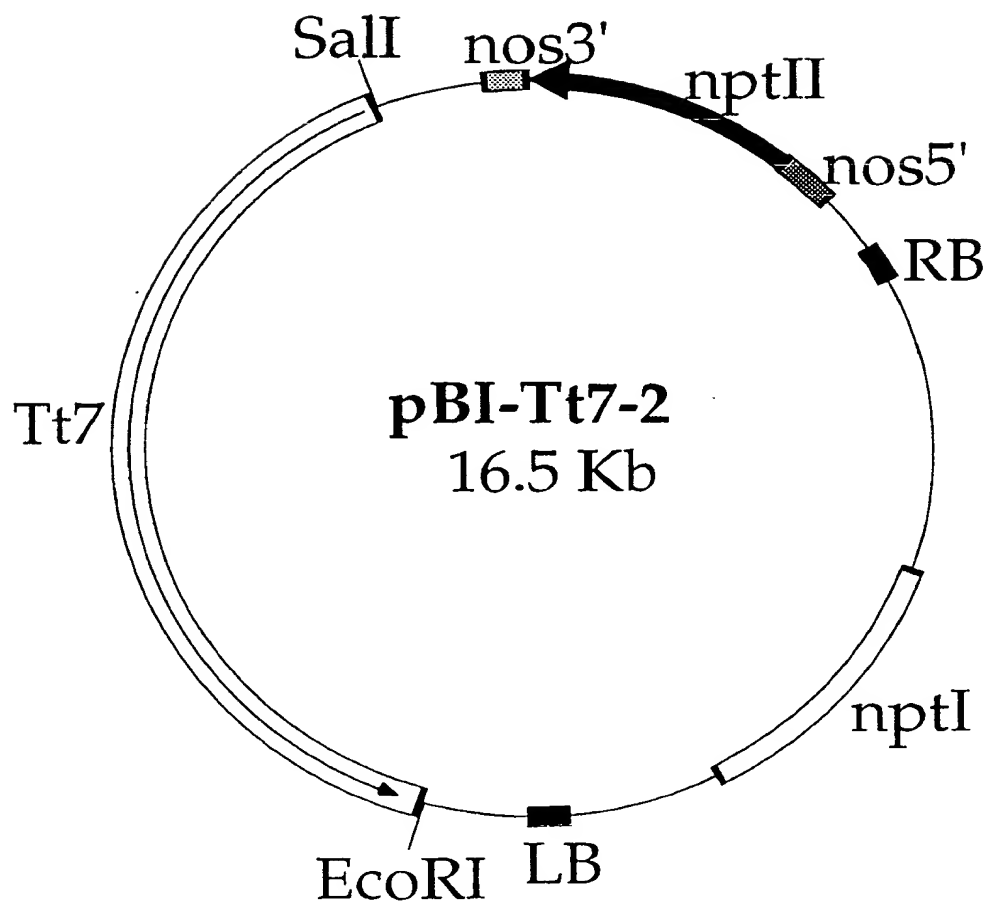


Figure 15

18/21

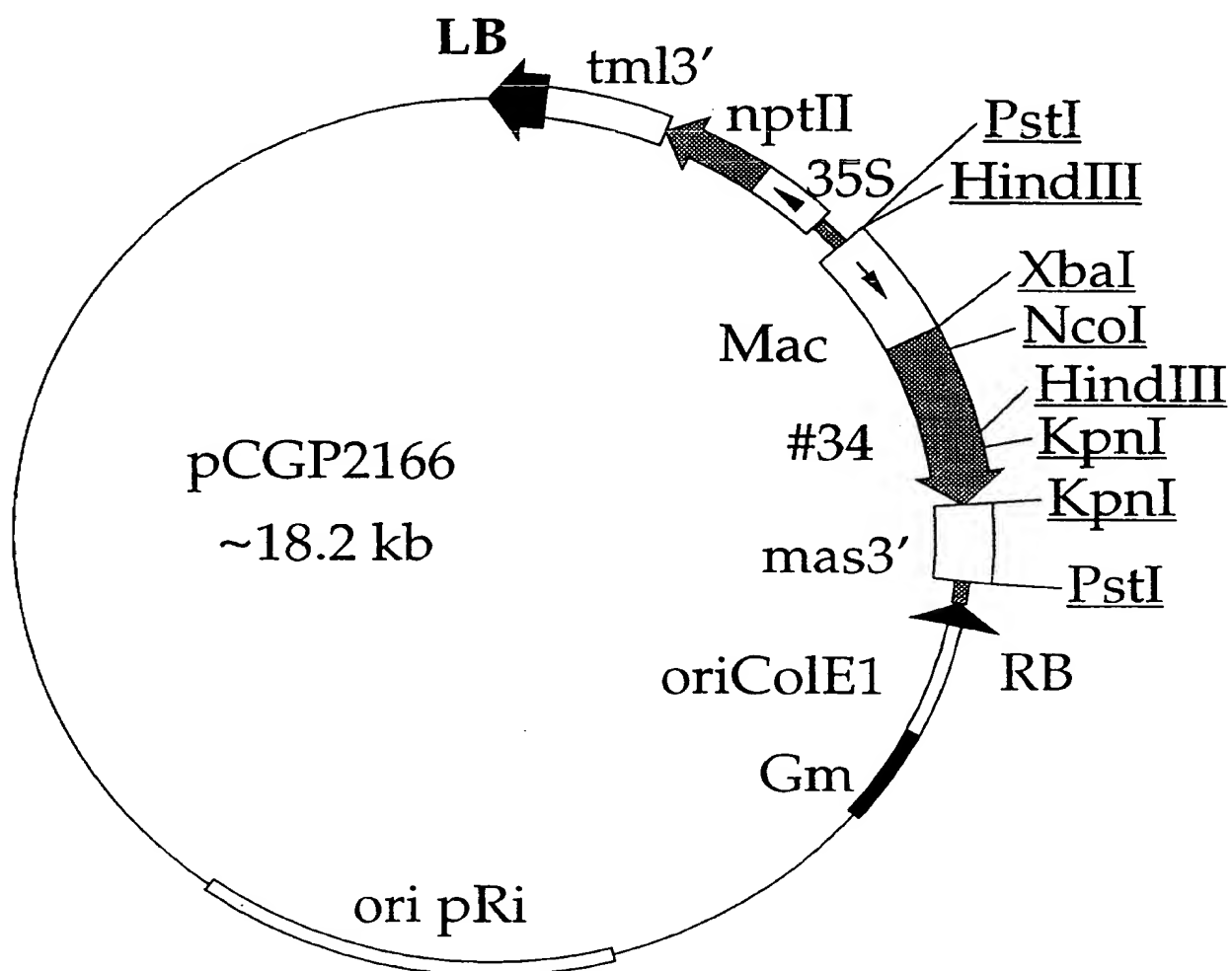


Figure 16

19/21

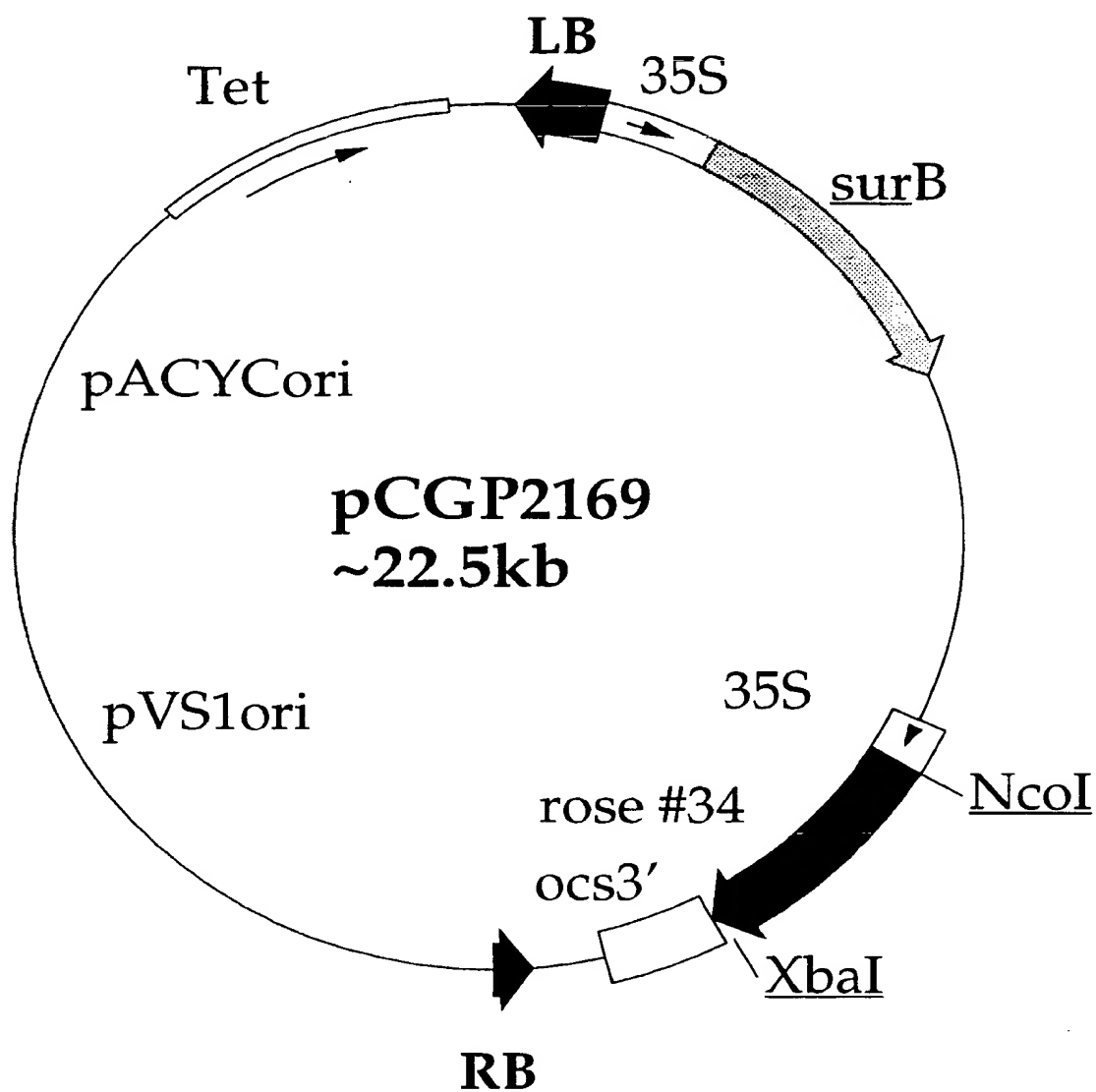


Figure 17

20/21

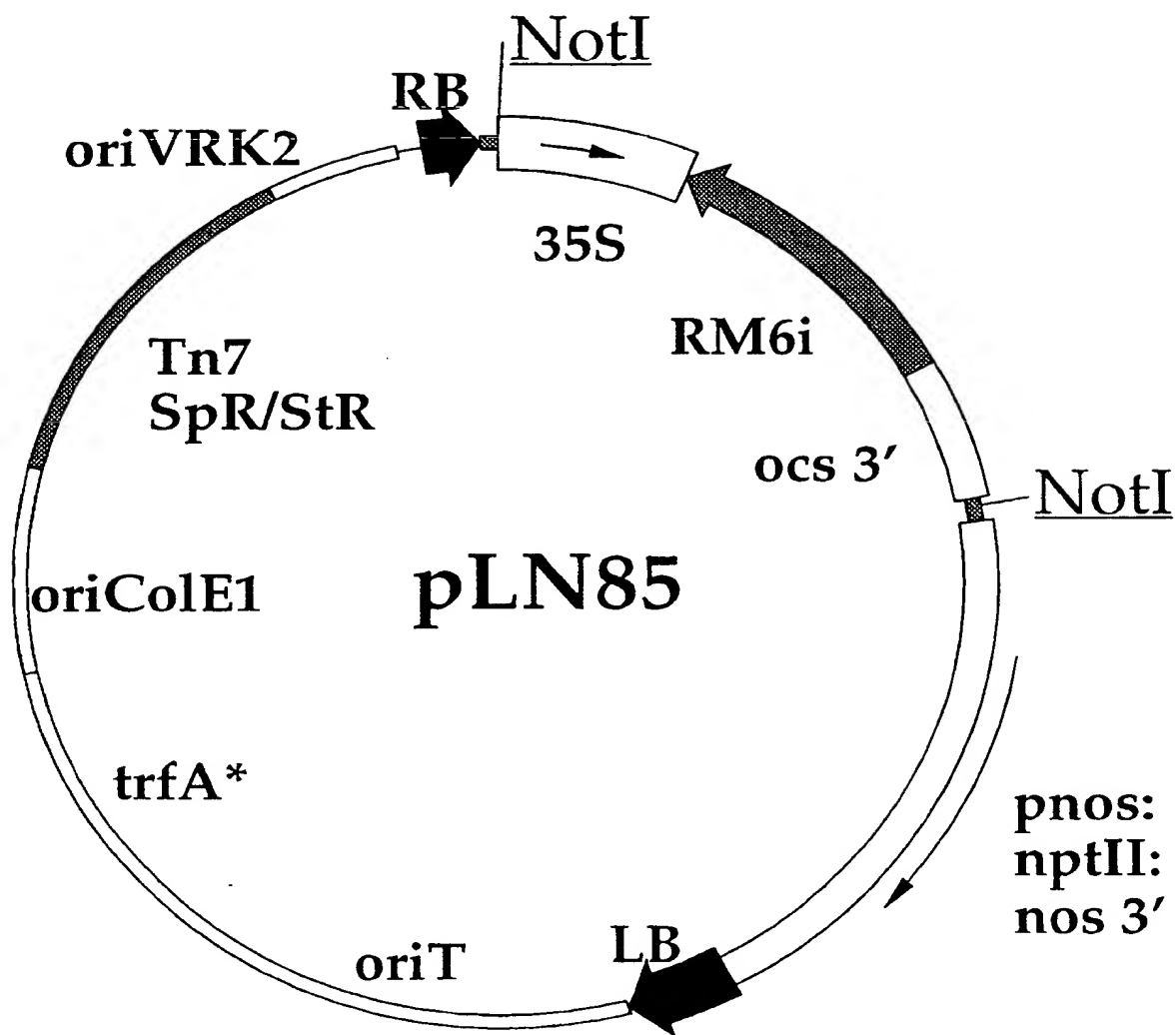


Figure 18

21/21

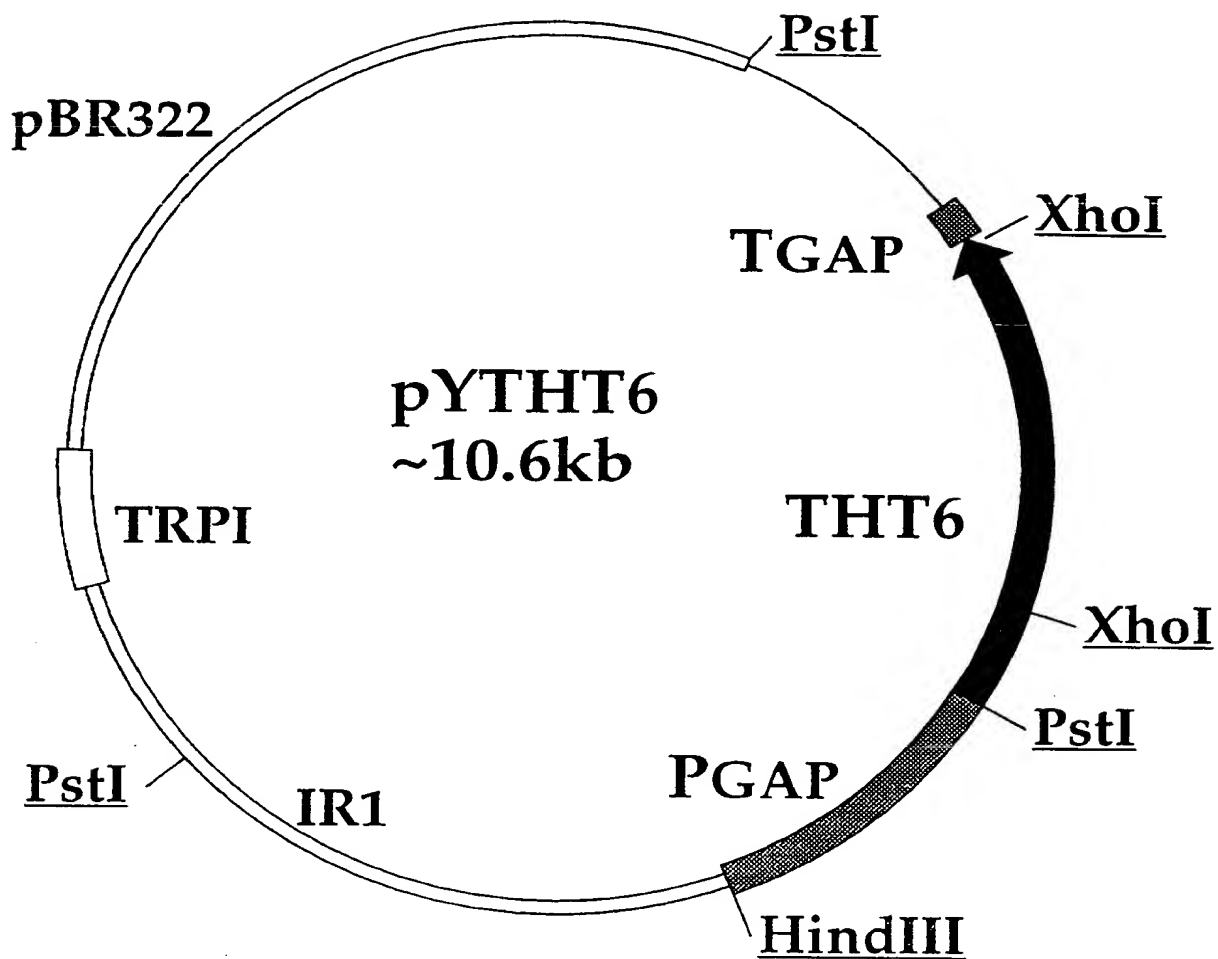


Figure 19

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/AU 97/00124**A. CLASSIFICATION OF SUBJECT MATTER**Int Cl⁶: C12N 15/29, 15/63
A01H 5/00, 5/02**B. FIELDS SEARCHED**Minimum documentation searched (classification system followed by classification symbols)
See electronic databases belowDocumentation searched other than minimum documentation to the extent that such documents are included in the fields searched
See electronic databases belowElectronic data base consulted during the international search (name of data base and, where practicable, search terms used)
WPAT, JAPIO, USPM - ss1: flavonoid(3n)monooxygenase# or flavonoid(3n)hydroxylase# or F3(O)H; ss2: A01H/IC or C12N-009/IC or C12N-015/IC; ss1 and ss2CHEMICAL ABSTRACTS - flavonoid 3' hydroxylase (clones pcgp772, pcgp602, pcgp619, pcgp635, pcgp854, pcgp773)/cn; flavonoid 3' hydroxylase/cn, flavonoid 3' monooxygenaseGENEBANK, EMBL, SWISS-PROT - Sequence searches**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	AU A 37413/93 (INTERNATIONAL FLOWER DEVELOPMENTS PTY. LTD.) 14 OCTOBER 1993 C12N 15/53, 15/11, 9/02, C12Q 1/66, A01H 5/00	27,28,30,31,38
A	PHYTOCHEMISTRY, VOL. 35 PAGES 145-150, 1994, SCHWINN KE et al. "Floral flavonoids and the potential for pelargonidin biosynthesis in commercial chrysanthemum cultivars"	1 - 38

☐ Further documents are listed in the continuation of Box C☒ See patent family annex

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
9 April 1997Date of mailing of the international search report
22 APR 1997Name and mailing address of the ISA/AU
AUSTRALIAN INDUSTRIAL PROPERTY ORGANISATION
PO BOX 200
WODEN ACT 2606
AUSTRALIA Facsimile No.: (06) 285 3929

Authorized officer

KAREN AYERS

Telephone No.: (06) 283 2082

INTERNATIONAL SEARCH REPORT
Information on patent family members

International Application No.
PCT/AU 97/00124

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report	Patent Family Member
AU, A, 37413/93	WO, A, 93/20206
	EP, A, 640136
	NZ, A, 249808
END OF ANNEX	